


MICROORGANISMS	
Optional Sheet in connection with the microorganism referred to on page 88, lines 1-12 of the description *	
<b>A. IDENTIFICATION OF DEPOSIT *</b> Further deposits are identified on an additional sheet *	
Name of depositary institution * American Type Culture Collection	
Address of depositary institution (including postal code and country) * 12301 Parklawn Drive Rockville, MD 20852 US	
Date of deposit * <u>May 2, 1991</u> Accession Number * <u>68610</u>	
<b>B. ADDITIONAL INDICATIONS *</b> (leave blank if not applicable). This information is continued on a separate attached sheet	
<b>C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *</b> (if the indications are not all designated States)	
<b>D. SEPARATE FURNISHING OF INDICATIONS *</b> (leave blank if not applicable) The indications listed below will be submitted to the International Bureau later * (Specify the general nature of the indications e.g., "Accession Number of Deposit")	
<b>E.</b> <input checked="" type="checkbox"/> This sheet was received with the International application when filed (to be checked by the receiving Office)	
<div style="text-align: center;"> (Authorized Officer)</div>	
<input type="checkbox"/> The date of receipt (from the applicant) by the International Bureau *	
was _____ (Authorized Officer)	

- 88.2 -

International Application No: PCT/ /

Form PCT/RO/134 (cont.)

American Type Culture Collection

12301 Parklawn Drive  
Rockville, MD 20852  
US

Accession No.

68609

68611

Date of Deposit

May 2, 1991

May 2, 1991

-89-

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Artavanis-Tsakonas, S. et al.
- (ii) TITLE OF INVENTION: Therapeutic And Diagnostic Methods  
And Compositions Based On Notch Proteins And  
Nucleic Acids
- (iii) NUMBER OF SEQUENCES: 21
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Pennie & Edmonds
  - (B) STREET: 1155 Avenue of the Americas
  - (C) CITY: New York
  - (D) STATE: New York
  - (E) COUNTRY: U.S.A.
  - (F) ZIP: 10036
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: To be assigned
  - (B) FILING DATE: On even date
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Misrock, S. Leslie
  - (B) REGISTRATION NUMBER: 18,872
  - (C) REFERENCE/DOCKET NUMBER: 7326-018
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 212 790-9090
  - (B) TELEFAX: 212 8698864/9741
  - (C) TELEX: 66141 PENNIE

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2892 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 142..2640

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GAATTCGGAG GAATTATTCA AACATAAAC ACAATAAACA ATTTGAGTAG TTGCCGCACA	60
CACACACACA CACAGCCCGT GGATTATTAC ACTAAAAGCG ACACTCAATC CAAAAAATCA	120
GCAACAAAAA CATCAATAAA C ATG CAT TGG ATT AAA TGT TTA TTA ACA GCA	171

-90-

Met His Trp Ile Lys Cys Leu Leu Thr Ala																		
1 5 10																		
TTC	ATT	TGC	TTC	ACA	GTC	ATC	GTG	CAG	GTT	CAC	AGT	TCC	GGC	AGC	TTT		219	
Phe	Ile	Cys	Phe	Thr	Val	Ile	Val	Gln	Val	His	Ser	Ser	Gly	Ser	Phe			
				15					20					25				
GAG	TTG	CGC	CTG	AAG	TAC	TTC	AGC	AAC	GAT	CAC	GGG	CGG	GAC	AAC	GAG		267	
Glu	Leu	Arg	Leu	Lys	Tyr	Phe	Ser	Asn	Asp	His	Gly	Arg	Asp	Asn	Glu			
				30					35					40				
GGT	CGC	TGC	TGC	AGC	GGG	GAG	TCG	GAC	GGA	GCG	ACG	GGC	AAG	TGC	CTG		315	
Gly	Arg	Cys	Cys	Ser	Gly	Glu	Ser	Asp	Gly	Ala	Thr	Gly	Lys	Cys	Leu			
				45					50					55				
GGC	AGC	TGC	AAG	ACG	CGG	TTT	CGC	GTC	TGC	CTA	AAG	CAC	TAC	CAG	GCC		363	
Gly	Ser	Cys	Lys	Thr	Arg	Phe	Arg	Val	Cys	Leu	Lys	His	Tyr	Gln	Ala			
				60					65					70				
ACC	ATC	GAC	ACC	ACC	TCC	CAG	TGC	ACC	TAC	GGG	GAC	GTG	ATC	ACG	CCC		411	
Thr	Ile	Asp	Thr	Thr	Ser	Gln	Cys	Thr	Tyr	Gly	Asp	Val	Ile	Thr	Pro			
				75					80					85				
ATT	CTC	GGC	GAG	AAC	TCG	GTC	AAT	CTG	ACC	GAC	GCC	CAG	CGC	TTC	CAG		459	
Ile	Leu	Gly	Glu	Asn	Ser	Val	Asn	Leu	Thr	Asp	Ala	Gln	Arg	Phe	Gln			
				95					100					105				
AAC	AAG	GGC	TTC	ACG	AAT	CCC	ATC	CAG	TTC	CCC	TTC	TCG	TTC	TCA	TGG		507	
Asn	Lys	Gly	Phe	Thr	Asn	Pro	Ile	Gln	Phe	Pro	Phe	Ser	Phe	Ser	Trp			
				110					115					120				
CCG	GGT	ACC	TTC	TCG	CTG	ATC	GTC	GAG	GCC	TGG	CAT	GAT	ACG	AAC	AAT		555	
Pro	Gly	Thr	Phe	Ser	Leu	Ile	Val	Glu	Ala	Trp	His	Asp	Thr	Asn	Asn			
				125					130					135				
AGC	GGC	AAT	GCG	CGA	ACC	AAC	AAG	CTC	CTC	ATC	CAG	CGA	CTC	TTG	GTG		603	
Ser	Gly	Asn	Ala	Arg	Thr	Asn	Lys	Leu	Leu	Ile	Gln	Arg	Leu	Leu	Val			
				140					145					150				
CAG	CAG	GTA	CTG	GAG	GTG	TCC	TCC	GAA	TGG	AAG	ACG	AAC	AAG	TCG	GAA		651	
Gln	Gln	Val	Leu	Glu	Val	Ser	Ser	Glu	Trp	Lys	Thr	Asn	Lys	Ser	Glu			
				155					160					165				
TCG	CAG	TAC	ACG	TCG	CTG	GAG	TAC	GAT	TTC	CGT	GTC	ACC	TGC	GAT	CTC		699	
Ser	Gln	Tyr	Thr	Ser	Leu	Glu	Tyr	Asp	Phe	Arg	Val	Thr	Cys	Asp	Leu			
				175					180					185				
AAC	TAC	TAC	GGA	TCC	GGC	TGT	GCC	AAG	TTC	TGC	CGG	CCC	CGC	GAC	GAT		747	
Asn	Tyr	Tyr	Gly	Ser	Gly	Cys	Ala	Lys	Phe	Cys	Arg	Pro	Arg	Asp	Asp			
				190					195					200				
TCA	TTT	GGA	CAC	TCG	ACT	TGC	TCG	GAG	ACG	GGC	GAA	ATT	ATC	TGT	TTG		795	
Ser	Phe	Gly	His	Ser	Thr	Cys	Ser	Glu	Thr	Gly	Glu	Ile	Ile	Cys	Leu			
				205					210					215				
ACC	GGA	TGG	CAG	GGC	GAT	TAC	TGT	CAC	ATA	CCC	AAA	TGC	GCC					



-91-

ATC CAT GGC ACC TGC AAC AAA CCC TGG ACT TGC ATC TGC AAC GAG GGT Ile His Gly Thr Cys Asn Lys Pro Trp Thr Cys Ile Cys Asn Glu Gly 270 275 280	987
TGG GGA GGC TTG TAC TGC AAC CAG GAT CTG AAC TAC TGC ACC AAC CAC Trp Gly Gly Leu Tyr Cys Asn Gln Asp Leu Asn Tyr Cys Thr Asn His 285 290 295	1035
AGA CCC TGC AAG AAT GGC GGA ACC TGC TTC AAC ACC GGC GAG GGA TTG Arg Pro Cys Lys Asn Gly Gly Thr Cys Phe Asn Thr Gly Glu Gly Leu 300 305 310	1083
TAC ACA TGC AAA TGC GCT CCA GGA TAC AGT GGT GAT GAT TGC GAA AAT Tyr Thr Cys Lys Cys Ala Pro Gly Tyr Ser Gly Asp Asp Cys Glu Asn 315 320 325 330	1131
GAG ATC TAC TCC TGC GAT GCC GAT GTC AAT CCC TGC CAG AAT GGT GGT Glu Ile Tyr Ser Cys Asp Ala Asp Val Asn Pro Cys Gln Asn Gly Gly 335 340 345	1179
ACC TGC ATC GAT GAG CCG CAC ACA AAA ACC GGC TAC AAG TGT CAT TGC Thr Cys Ile Asp Glu Pro His Thr Lys Thr Gly Tyr Lys Cys His Cys 350 355 360	1227
GCC AAC GGC TGG AGC GGA AAG ATG TGC GAG GAG AAA GTG CTC ACG TGT Ala Asn Gly Trp Ser Gly Lys Met Cys Glu Glu Lys Val Leu Thr Cys 365 370 375	1275
TCG GAC AAA CCC TGT CAT CAG GGA ATC TGC CGC AAC GTT CGT CCT GGC Ser Asp Lys Pro Cys His Gln Gly Ile Cys Arg Asn Val Arg Pro Gly 380 385 390	1323
TTG GGA AGC AAG GGT CAG GGC TAC CAG TGC GAA TGT CCC ATT GGC TAC Leu Gly Ser Lys Gly Gln Gly Tyr Gln Cys Glu Cys Pro Ile Gly Tyr 395 400 405 410	1371
AGC GGA CCC AAC TGC GAT CTC CAG CTG GAC AAC TGC AGT CCG AAT CCA Ser Gly Pro Asn Cys Asp Leu Gln Leu Asp Asn Cys Ser Pro Asn Pro 415 420 425	1419
TGC ATA AAC GGT GGA AGC TGT CAG CCG AGC GGA AAG TGT ATT TGC CCA Cys Ile Asn Gly Gly Ser Cys Gln Pro Ser Gly Lys Cys Ile Cys Pro 430 435 440	1467
GCG GGA TTT TCG GGA ACG AGA TGC GAG ACC AAC ATT GAC GAT TGT CTT Ala Gly Phe Ser Gly Thr Arg Cys Glu Thr Asn Ile Asp Asp Cys Leu 445 450 455	1515
GGC CAC CAG TGC GAG AAC GGA GGC ACC TGC ATA GAT ATG GTC AAC CAA Gly His Gln Cys Glu Asn Gly Gly Thr Cys Ile Asp Met Val Asn Gln 460 465 470	1563
TAT CGC TGC CAA TGC GTT CCC GGT TTC CAT GGC ACC CAC TGT AGT AGC Tyr Arg Cys Gln Cys Val Pro Gly Phe His Gly Thr His Cys Ser Ser 475 480 485 490	1611
AAA GTT GAC TTG TGC CTC ATC AGA CCG TGT GCC AAT GGA GGA ACC TGC Lys Val Asp Leu Cys Leu Ile Arg Pro Cys Ala Asn Gly Gly Thr Cys 495 500 505	1659
TTG AAT CTC AAC AAC GAT TAC CAG TGC ACC TGT CGT GCG GGA TTT ACT Leu Asn Leu Asn Asn Asp Tyr Gln Cys Thr Cys Arg Ala Gly Phe Thr 510 515 520	1707
GGC AAG GAT TGC TCT GTG GAC ATC GAT GAG TGC AGC AGT GGA CCC TGT Gly Lys Asp Cys Ser Val Asp Ile Asp Glu Cys Ser Ser Gly Pro Cys 525 530 535	1755

-92-

CAT His 540	AAC Asn 540	GGC Gly 540	GGC Gly 540	ACT Thr 540	TGC Cys 545	ATG Met 545	AAC Asn 545	CGC Arg 545	GTC Val 545	AAT Asn 550	TCG Ser 550	TTC Phe 550	GAA Glu 550	TGC Cys 550	GTG Val 550	1803
TGT Cys 555	GCC Ala 555	AAT Asn 555	GGT Gly 555	TTC Phe 560	AGG Arg 560	GGC Gly 560	AAG Lys 560	CAG Gln 560	TGC Cys 565	GAT Asp 565	GAG Glu 565	GAG Glu 565	TCC Ser 565	TAC Tyr 570	GAT Asp 570	1851
TCG Ser 575	GTG Val 575	ACC Thr 575	TTC Phe 575	GAT Asp 575	GCC Ala 575	CAC His 575	CAA Gln 575	TAT Tyr 580	GGA Gly 580	GCG Ala 580	ACC Thr 580	ACA Thr 585	CAA Gln 585	GCG Ala 585	AGA Arg 585	1899
GCC Ala 590	GAT Asp 590	GGT Gly 590	TTG Leu 590	ACC Thr 590	AAT Asn 590	GCC Ala 595	CAG Gln 595	GTA Val 595	GTC Val 595	CTA Leu 595	ATT Ile 595	GCT Ala 600	GTT Val 600	TTC Phe 600	TCC Ser 600	1947
GTT Val 605	GCG Ala 605	ATG Met 605	CCT Pro 605	TTG Leu 605	GTG Val 605	GCG Ala 610	GTT Val 610	ATT Ile 610	GCG Ala 610	GCG Ala 610	TGC Cys 615	GTG Val 615	GTC Val 615	TTC Phe 615	TGC Cys 615	1995
ATG Met 620	AAG Lys 620	CGC Arg 620	AAG Lys 620	CGT Arg 625	AAG Lys 625	CGT Arg 625	GCT Ala 625	CAG Gln 625	GAA Glu 625	AAG Lys 630	GAC Asp 630	GAC Asp 630	GCG Ala 630	GAG Glu 630	GCC Ala 630	2043
AGG Arg 635	AAG Lys 635	CAG Gln 635	AAC Asn 635	GAA Glu 640	CAG Gln 640	AAT Asn 640	GCG Ala 640	GTG Val 640	GCC Ala 645	ACA Thr 645	ATG Met 645	CAT His 645	CAC His 645	AAT Asn 650	GGC Gly 650	2091
AGT Ser 655	GGG Gly 655	GTG Val 655	GGT Gly 655	GTA Val 655	GCT Ala 655	TTG Leu 655	GCT Ala 660	TCA Ser 660	GCC Ala 660	TCT Ser 660	CTG Leu 660	GGC Gly 665	GGC Gly 665	AAA Lys 665	ACT Thr 665	2139
GGC Gly 670	AGC Ser 670	AAC Asn 670	AGC Ser 670	GGT Gly 670	CTC Leu 675	ACC Thr 675	TTC Phe 675	GAT Asp 675	GGC Gly 675	GGC Gly 675	AAC Asn 680	CCG Pro 680	AAT Asn 680	ATC Ile 680	ATC Ile 680	2187
AAA Lys 685	AAC Asn 685	ACC Thr 685	TGG Trp 685	GAC Asp 685	AAG Lys 690	TCG Ser 690	GTC Val 690	AAC Asn 690	AAC Asn 690	ATT Ile 695	TGT Cys 695	GCC Ala 695	TCA Ser 695	GCA Ala 695	GCA Ala 695	2235
GCA Ala 700	GCG Ala 700	GCG Ala 700	GCG Ala 700	GCA Ala 705	GCA Ala 705	GCG Ala 705	GCG Ala 705	GAC Asp 705	GAG Glu 710	TGT Cys 710	CTC Leu 710	ATG Met 710	TAC Tyr 710	GGC Gly 710		2283
GGA Gly 715	TAT Tyr 715	GTG Val 715	GCC Ala 715	TCG Ser 720	GTG Val 720	GCG Ala 720	GAT Asp 720	AAC Asn 725	AAC Asn 725	AAT Asn 725	GCC Ala 725	AAC Asn 725	TCA Ser 725	GAC Asp 730	TTT Phe 730	2331
TGT Cys 735	GTG Val 735	GCT Ala 735	CCG Pro 735	CTA Leu 735	CAA Gln 735	AGA Arg 735	GCC Ala 740	AAG Lys 740	TCG Ser 740	CAA Gln 740	AAG Lys 740	CAA Gln 745	CTC Leu 745	AAC Asn 745	ACC Thr 745	2379
GAT Asp 750	CCC Pro 750	ACG Thr 750	CTC Leu 750	ATG Met 750	CAC His 755	CGC Arg 755	GGT Gly 755	TCG Ser 755	CCG Pro 755	GCA Ala 760	GGC Gly 760	AGC Ser 760	TCA Ser 760	GCC Ala 760	AAG Lys 760	2427
GGA Gly 765	GCG Ala 765	TCT Ser 765	GGC Gly 765	GGA Gly 770	GGA Gly 770	CCG Pro 770	GGA Gly 770	GCG Ala 770	GCG Ala 770	GAG Glu 775	GGC Gly 775	AAG Lys 775	AGG Arg 775	ATC Ile 775	TCT Ser 775	2475
GTT Val 780	TTA Leu 780	GGC Gly 780	GAG Glu 780	GGT Gly 785	TCC Ser 785	TAC Tyr 785	TGT Cys 785	AGC Ser 785	CAG Gln 785	CGT Arg 785	TGG Trp 790	CCC Pro 790	TCG Ser 790	TTG Leu 790	GCG Ala 790	2523
GCG Ala 795	GCG Ala 795	GGA Gly 795	GTG Val 795	GCC Ala 800	GGA Gly 800	GCC Ala 800	TGT Cys 805	TCA Ser 805	TCC Ser 805	CAG Gln 805	CTA Leu 805	ATG Met 810	GCT Ala 810	GCA Ala 810	GCT Ala 810	2571

-93-

TCG GCA GCG GGC AGC GGA GCG GGG ACG GCG CAA CAG CAG CGA TCC GTG 2619  
 Ser Ala Ala Gly Ser Gly Ala Gly Thr Ala Gln Gln Gln Arg Ser Val  
                   815                                  820                                  825

GTC TGC GGC ACT CCG CAT ATG TAACTCCAAA AATCCGGAAG GGCTCCTGGT 2670  
 Val Cys Gly Thr Pro His Met  
                   830

AAATCCGGAG AAATCCGCAT GGAGGAGCTG ACAGCACATA CACAAAGAAA AGACTGGGTT 2730

GGGTTCAAAA TGTGAGAGAG ACGCCAAAAT GTTGTGTGTG ATTGAAGCAG TTTAGTCGTC 2790

ACGAAAAATG AAAAATCTGT AACAGGCATA ACTCGTAAAC TCCCTAAAAA ATTTGTATAG 2850

TAATTAGCAA AGCTGTGACC CAGCCGTTTC GATCCCGAAT TC 2892

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 833 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met His Trp Ile Lys Cys Leu Leu Thr Ala Phe Ile Cys Phe Thr Val  
   1                                  5                                  10                                  15

Ile Val Gln Val His Ser Ser Gly Ser Phe Glu Leu Arg Leu Lys Tyr  
                   20                                  25                                  30

Phe Ser Asn Asp His Gly Arg Asp Asn Glu Gly Arg Cys Cys Ser Gly  
                   35                                  40                                  45

Glu Ser Asp Gly Ala Thr Gly Lys Cys Leu Gly Ser Cys Lys Thr Arg  
                   50                                  55                                  60

Phe Arg Val Cys Leu Lys His Tyr Gln Ala Thr Ile Asp Thr Thr Ser  
                   65                                  70                                  75                                  80

Gln Cys Thr Tyr Gly Asp Val Ile Thr Pro Ile Leu Gly Glu Asn Ser  
                   85                                  90                                  95

Val Asn Leu Thr Asp Ala Gln Arg Phe Gln Asn Lys Gly Phe Thr Asn  
                   100                                  105                                  110

Pro Ile Gln Phe Pro Phe Ser Phe Ser Trp Pro Gly Thr Phe Ser Leu  
                   115                                  120                                  125

Ile Val Glu Ala Trp His Asp Thr Asn Asn Ser Gly Asn Ala Arg Thr  
                   130                                  135                                  140

Asn Lys Leu Leu Ile Gln Arg Leu Leu Val Gln Gln Val Leu Glu Val  
                   145                                  150                                  155                                  160

Ser Ser Glu Trp Lys Thr Asn Lys Ser Glu Ser Gln Tyr Thr Ser Leu  
                   165                                  170                                  175

Glu Tyr Asp Phe Arg Val Thr Cys Asp Leu Asn Tyr Tyr Gly Ser Gly  
                   180                                  185                                  190

Cys Ala Lys Phe Cys Arg Pro Arg Asp Asp Ser Phe Gly His Ser Thr  
                   195                                  200                                  205

-94-

Cys Ser Glu Thr Gly Glu Ile Ile Cys Leu Thr Gly Trp Gln Gly Asp  
 210 215 220  
 Tyr Cys His Ile Pro Lys Cys Ala Lys Gly Cys Glu His Gly His Cys  
 225 230 235 240  
 Asp Lys Pro Asn Gln Cys Val Cys Gln Leu Gly Trp Lys Gly Ala Leu  
 245 250 255  
 Cys Asn Glu Cys Val Leu Glu Pro Asn Cys Ile His Gly Thr Cys Asn  
 260 265 270  
 Lys Pro Trp Thr Cys Ile Cys Asn Glu Gly Trp Gly Gly Leu Tyr Cys  
 275 280 285  
 Asn Gln Asp Leu Asn Tyr Cys Thr Asn His Arg Pro Cys Lys Asn Gly  
 290 295 300  
 Gly Thr Cys Phe Asn Thr Gly Glu Gly Leu Tyr Thr Cys Lys Cys Ala  
 305 310 315 320  
 Pro Gly Tyr Ser Gly Asp Asp Cys Glu Asn Glu Ile Tyr Ser Cys Asp  
 325 330 335  
 Ala Asp Val Asn Pro Cys Gln Asn Gly Gly Thr Cys Ile Asp Glu Pro  
 340 345 350  
 His Thr Lys Thr Gly Tyr Lys Cys His Cys Ala Asn Gly Trp Ser Gly  
 355 360 365  
 Lys Met Cys Glu Glu Lys Val Leu Thr Cys Ser Asp Lys Pro Cys His  
 370 375 380  
 Gln Gly Ile Cys Arg Asn Val Arg Pro Gly Leu Gly Ser Lys Gly Gln  
 385 390 395 400  
 Gly Tyr Gln Cys Glu Cys Pro Ile Gly Tyr Ser Gly Pro Asn Cys Asp  
 405 410 415  
 Leu Gln Leu Asp Asn Cys Ser Pro Asn Pro Cys Ile Asn Gly Gly Ser  
 420 425 430  
 Cys Gln Pro Ser Gly Lys Cys Ile Cys Pro Ala Gly Phe Ser Gly Thr  
 435 440 445  
 Arg Cys Glu Thr Asn Ile Asp Asp Cys Leu Gly His Gln Cys Glu Asn  
 450 455 460  
 Gly Gly Thr Cys Ile Asp Met Val Asn Gln Tyr Arg Cys Gln Cys Val  
 465 470 475 480  
 Pro Gly Phe His Gly Thr His Cys Ser Ser Lys Val Asp Leu Cys Leu  
 485 490 495  
 Ile Arg Pro Cys Ala Asn Gly Gly Thr Cys Leu Asn Leu Asn Asn Asp  
 500 505 510  
 Tyr Gln Cys Thr Cys Arg Ala Gly Phe Thr Gly Lys Asp Cys Ser Val  
 515 520 525  
 Asp Ile Asp Glu Cys Ser Ser Gly Pro Cys His Asn Gly Gly Thr Cys  
 530 535 540  
 Met Asn Arg Val Asn Ser Phe Glu Cys Val Cys Ala Asn Gly Phe Arg  
 545 550 555 560  
 Gly Lys Gln Cys Asp Glu Glu Ser Tyr Asp Ser Val Thr Phe Asp Ala

565

570

575

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1320 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

```
(ix) FEATURE:
      (A) NAME/KEY: CDS
      (B) LOCATION: 442..1320
```

-96-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCGAGTCGAG CGCCGTGCTT CGAGCGGTGA TGAGCCCCTT TTCTGTCAAC GCTAAAGATC	60
TACAAAACAT CAGCGCCTAT CAAGTGAAG TGTCAAGTGT GAACAAAACA AAAACGAGAG	120
AAGCACATAC TAAGGTCCAT ATAAATAATA AATAATAATT GTGTGTGATA ACAACATTAT	180
CCAAACAAAA CCAAACAAAA CGAAGGCAAA GTGGAGAAAA TGATACAGCA TCCAGAGTAC	240
GGCCGTTATT CAGCTATCCA GAGCAAGTGT AGTGTGGCAA AATAGAAACA AACAAAGGCA	300
CCAAAATCTG CATACATGGG CTAATTAAGG CTGCCCAGCG AATTTACATT TGTGTGGTGC	360
CAATCCAGAG TGAATCCGAA ACAAACTCCA TCTAGATCGC CAACCAGCAT CACGCTCGCA	420
AACGCCCCCA GAATGTACAA A ATG TTT AGG AAA CAT TTT CGG CGA AAA CCA	471
Met Phe Arg Lys His Phe Arg Arg Lys Pro	10
1 5	
GCT ACG TCG TCG TCG TTG GAG TCA ACA ATA GAA TCA GCA GAC AGC CTG	519
Ala Thr Ser Ser Ser Leu Glu Ser Thr Ile Glu Ser Ala Asp Ser Leu	25
15 20	
GGA ATG TCC AAG AAG ACG GCG ACA AAA AGG CAG CGT CCG AGG CAT CGG	567
Gly Met Ser Lys Lys Thr Ala Thr Lys Arg Gln Arg Pro Arg His Arg	40
30 35	
GTA CCC AAA ATC GCG ACC CTG CCA TCG ACG ATC CGC GAT TGT CGA TCA	615
Val Pro Lys Ile Ala Thr Leu Pro Ser Thr Ile Arg Asp Cys Arg Ser	55
45 50	
TTA AAG TCT GCC TGC AAC TTA ATT GCT TTA ATT TTA ATA CTG TTA GTC	663
Leu Lys Ser Ala Cys Asn Leu Ile Ala Leu Ile Leu Ile Leu Leu Val	70
60 65	
CAT AAG ATA TCC GCA GCT GGT AAC TTC GAG CTG GAA ATA TTA GAA ATC	711
His Lys Ile Ser Ala Ala Gly Asn Phe Glu Leu Glu Ile Leu Glu Ile	85 90
75 80	
TCA AAT ACC AAC AGC CAT CTA CTC AAC GGC TAT TGC TGC GGC ATG CCA	759
Ser Asn Thr Asn Ser His Leu Leu Asn Gly Tyr Cys Cys Gly Met Pro	100 105
95 100	
GCG GAA CTT AGG GCC ACC AAG ACG ATA GGC TGC TCG CCA TGC ACG ACG	807
Ala Glu Leu Arg Ala Thr Lys Thr Ile Gly Cys Ser Pro Cys Thr Thr	110 115 120
110 115	
GCA TTC CGG CTG TGC CTG AAG GAG TAC CAG ACC ACG GAG CAG GGT GCC	855
Ala Phe Arg Leu Cys Leu Lys Glu Tyr Gln Thr Thr Glu Gln Gly Ala	125 130 135
125 130	
AGC ATA TCC ACG GGC TGT TCG TTT GGC AAC GCC ACC ACC AAG ATA CTG	903
Ser Ile Ser Thr Gly Cys Ser Phe Gly Asn Ala Thr Thr Lys Ile Leu	140 145 150
140 145	
GGT GGC TCC AGC TTT GTG CTC AGC GAT CCG GGT GTG GGA GCC ATT GTG	951
Gly Gly Ser S r Phe Val Leu Ser Asp Pro Gly Val Gly Ala Ile Val	155 160 165 170
155 160	
CTG CCC TTT ACG TTT CGT TGG ACG AAG TCG TTT ACG CTG ATA CTG CAG	999
Leu Pro Phe Thr Phe Arg Trp Thr Lys Ser Phe Thr Leu Il Leu Gln	175 180 185
175 180	
GCG TTG GAT ATG TAC AAC ACA TCC TAT CCA GAT GCG GAG AGG TTA ATT	1047

-97-

Ala	Leu	Asp	Met	Tyr	Asn	Thr	Ser	Tyr	Pro	Asp	Ala	Glu	Arg	Leu	Ile		
			190					195					200				
GAG	GAA	ACA	TCA	TAC	TCG	GGC	GTG	ATA	CTG	CCG	TCG	CCG	GAG	TGG	AAG		1095
Glu	Glu	Thr	Ser	Tyr	Ser	Gly	Val	Ile	Leu	Pro	Ser	Pro	Glu	Trp	Lys		
		205					210					215					
ACG	CTG	GAC	CAC	ATC	GGG	CGG	AAC	GCG	CGG	ATC	ACC	TAC	CGT	GTC	CGG		1143
Thr	Leu	Asp	His	Ile	Gly	Arg	Asn	Ala	Arg	Ile	Thr	Tyr	Arg	Val	Arg		
	220					225					230						
GTG	CAA	TGC	GCC	GTT	ACC	TAC	TAC	AAC	ACG	ACC	TGC	ACG	ACC	TTC	TGC		1191
Val	Gln	Cys	Ala	Val	Thr	Tyr	Tyr	Asn	Thr	Thr	Cys	Thr	Thr	Phe	Cys		
235					240					245					250		
CGT	CCG	CGG	GAC	GAT	CAG	TTC	GGT	CAC	TAC	GCC	TGC	GGC	TCC	GAG	GGT		1239
Arg	Pro	Arg	Asp	Asp	Gln	Phe	Gly	His	Tyr	Ala	Cys	Gly	Ser	Glu	Gly		
			255						260					265			
CAG	AAG	CTC	TGC	CTG	AAT	GGC	TGG	CAG	GGC	GTC	AAC	TGC	GAG	GAG	GCC		1287
Gln	Lys	Leu	Cys	Leu	Asn	Gly	Trp	Gln	Gly	Val	Asn	Cys	Glu	Glu	Ala		
			270					275					280				
ATA	TGC	AAG	GCG	GGC	TGC	GAC	CCC	GTC	CAC	GGC							1320
Ile	Cys	Lys	Ala	Gly	Cys	Asp	Pro	Val	His	Gly							
		285					290										

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 293 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Phe	Arg	Lys	His	Phe	Arg	Arg	Lys	Pro	Ala	Thr	Ser	Ser	Ser	Leu		
1				5					10					15			
Glu	Ser	Thr	Ile	Glu	Ser	Ala	Asp	Ser	Leu	Gly	Met	Ser	Lys	Lys	Thr		
		20					25						30				
Ala	Thr	Lys	Arg	Gln	Arg	Pro	Arg	His	Arg	Val	Pro	Lys	Ile	Ala	Thr		
		35				40					45						
Leu	Pro	Ser	Thr	Ile	Arg	Asp	Cys	Arg	Ser	Leu	Lys	Ser	Ala	Cys	Asn		
	50				55					60							
Leu	Ile	Ala	Leu	Ile	Leu	Ile	Leu	Leu	Val	His	Lys	Ile	Ser	Ala	Ala		
65				70					75					80			
Gly	Asn	Phe	Glu	Leu	Glu	Ile	Leu	Glu	Ile	Ser	Asn	Thr	Asn	Ser	His		
			85					90					95				
Leu	Leu	Asn	Gly	Tyr	Cys	Cys	Gly	Met	Pro	Ala	Glu	Leu	Arg	Ala	Thr		
		100					105					110					
Lys	Thr	Ile	Gly	Cys	Ser	Pro	Cys	Thr	Thr	Ala	Phe	Arg	Leu	Cys	Leu		
	115					120					125						
Lys	Glu	Tyr	Gln	Thr	Thr	Glu	Gln	Gly	Ala	Ser	Ile	Ser	Thr	Gly	Cys		
	130					135					140						
Ser	Phe	Gly	Asn	Ala	Thr	Thr	Lys	Ile	Leu	Gly	Gly	Ser	Ser	Phe	Val		

-98-

145		150		155		160									
Leu	Ser	Asp	Pro	Gly	Val	Gly	Ala	Ile	Val	Leu	Pro	Phe	Thr	Phe	Arg
				165					170					175	
Trp	Thr	Lys	Ser	Phe	Thr	Leu	Ile	Leu	Gln	Ala	Leu	Asp	Met	Tyr	Asn
			180					185					190		
Thr	Ser	Tyr	Pro	Asp	Ala	Glu	Arg	Leu	Ile	Glu	Glu	Thr	Ser	Tyr	Ser
		195					200					205			
Gly	Val	Ile	Leu	Pro	Ser	Pro	Glu	Trp	Lys	Thr	Leu	Asp	His	Ile	Gly
	210					215					220				
Arg	Asn	Ala	Arg	Ile	Thr	Tyr	Arg	Val	Arg	Val	Gln	Cys	Ala	Val	Thr
225					230					235					240
Tyr	Tyr	Asn	Thr	Thr	Cys	Thr	Thr	Phe	Cys	Arg	Pro	Arg	Asp	Asp	Gln
				245					250					255	
Phe	Gly	His	Tyr	Ala	Cys	Gly	Ser	Glu	Gly	Gln	Lys	Leu	Cys	Leu	Asn
			260					265					270		
Gly	Trp	Gln	Gly	Val	Asn	Cys	Glu	Glu	Ala	Ile	Cys	Lys	Ala	Gly	Cys
		275					280					285			
Asp	Pro	Val	His	Gly											
				290											

## (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 267 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGGTGGACTT CCTTCGTGTA TTGGTGGGAG CCCTCGGGAA CGGGGGGTAA CACTGAAAGG	60
TCGAGTACCC ATTTCCGTCA TAACGGGTTG GTCGCCCCCT AGGGGTCGGA GTCAGGTGGA	120
CGGGAGGTCG ACAACGCCCCG GGGGACGGGT GGTACATGTT GTAAGGTCTT TACCGGACCG	180
GGCAAACGGG TCACACCGAA AGGGGTGAAC GGTAACACG GGGTCGTCCT GCCCGTCCAT	240
CGAGTCTGGT AAGAGGGTCG CCTTAAG	267

## (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 574 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA



-99-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GAATTCCTTC CATTATACGT GACTTTTCTG AAACGTAGC CACCCTAGTG TCTCTAACTC	60
CCTCTGGAGT TTGTCAGCTT TGGTCTTTTC AAAGAGCAGG CTCTCTTCAA GCTCCTTAAT	120
GCGGGCATGC TCCAGTTTGG TCTGCGTCTC AAGATCACCT TTGGTAATTG ATTCTTCTTC	180
AACCCGGAAC TGAAGGCTGG CTCTCACCT CTAGGCAGAG CAGGAATTCC GAGGTGGATG	240
TGTTAGATGT GAATGTCCGT GGCCAGATG GCTGCACCCC ATTGATGTTG GCTTCTCTCC	300
GAGGAGGCAG CTCAGATTG AGTGATGAAG ATGAAGATGC AGAGGACTGT TCTGCTAACA	360
TCATCACAGA CTTGGTCTAC CAGGGTGCCA GCCTCCAGNC CAGACAGACC GGACTGGTGA	420
GATGGCCCTG CACCTTGCG CCGCTACTC ACGGGCTGAT GCTGCCAAGC GTCTCCTGGA	480
TGCAGGTGCA GATGCCAATG CCCAGGACAA CATGGGCCGC TGTCCACTCC ATGCTGCAGT	540
GGCACGTGAT GCCAAGGTGT ATTCAGATCT GTTA	574

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 295 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TCCAGATTCT GATTGCAAC CGAGTAACTG ATCTAGATGC CAGGATGAAT GATGGTACTA	60
CACCCCTGAT CCTGGCTGCC CGCCTGGCTG TGGAGGAAT GGTGGCAGAA CTGATCAACT	120
GCCAAGCGGA TGTGAATGCA GTGGATGACC ATGGAAAATC TGCTCTTCAC TGGGCAGCTG	180
CTGTCAATAA TGTGGAGGCA ACTCTTTTGT TGTGAAAAA TGGGGCCAAC CGAGACATGC	240
AGGACAACAA GGAAGAGACA CCTCTGTTTC TTGCTGCCCG GGAGGAGCTA TAAGC	295

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 248 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GAATTCCATT CAGGAGGAAA GGGTGGGGAG AGAAGCAGGC ACCCACTTTC CCGTGGCTGG	60
ACTCGTTCCC AGGTGGCTCC ACCGGCAGCT GTGACCGCCG CAGGTGGGGG CGGAGTGCCA	120
TTCAGAAAAT TCCAGAAAAG CCCTACCCCA ACTCGGACGG CAACGTCACA CCCGTGGGTA	180

-100-

GCAACTGGCA CACAAACAGC CAGCGTGTCT GGGGCACGGG GGGATGGCAC CCCCTGCAGG 240  
CAGAGCTG 248

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 323 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TACGTATCTC GAGCACAGAC AGCTGACGTA CACTTTTNNNA GTGCGAGGGA CATTGTCGG 60  
 ACCAGTACGA ACATTTAGGC TCAGTACGGT AGGTCCATGG CCAAGACTAG GAGACGTAGG 120  
 GAGCTACAGG TCCCGCTCGC TAAACTCGGA CCACTGAAAC CTCCGGTCGA CAGTCGGTAA 180  
 GCGAACAAGA GGGCCAGATC TTAGAGAAGG TGTCGCGGCG AGACTCGGGC TCGGGTCAGG 240  
 CGGCCTTAAG GACGTCGGGC CCNNNAGGTG ATCAAGATCT CGNCNCGGCG GGCGCCACCT 300  
 CGAGGNCGAA AACAAGGGAA ATC 323

## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3234 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: CDS  
 (B) LOCATION: 1..3234

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TGC CAG GAG GAC GCG GGC AAC AAG GTC TGC AGC CTG CAG TGC AAC AAC 48  
 Cys Gln Glu Asp Ala Gly Asn Lys Val Cys Ser Leu Gln Cys Asn Asn 15  
 1 5 10  
 CAC GCG TGC GGC TGG GAC GGC GGT GAC TGC TCC CTC AAC TTC AAT GAC 96  
 His Ala Cys Gly Trp Asp Gly Gly Asp Cys Ser Leu Asn Phe Asn Asp 20 25 30  
 CCC TGG AAG AAC TGC ACG CAG TCT CTG CAG TGC TGG AAG TAC TTC AGT 144  
 Pro Trp Lys Asn Cys Thr Gln Ser Leu Gln Cys Trp Lys Tyr Phe Ser 35 40 45  
 GAC GGC CAC TGT GAC AGC CAG TGC AAC TCA GCC GGC TGC CTC TTC GAC 192  
 Asp Gly His Cys Asp Ser Gln Cys Asn Ser Ala Gly Cys Leu Phe Asp 50 55 60  
 GGC TTT GAC TGC CAG CGT GCG GAA GGC CAG TGC AAC CCC CTG TAC GAC 240

-101-

Gly 65	Phe	Asp	Cys	Gln	Arg 70	Ala	Glu	Gly	Gln	Cys 75	Asn	Pro	Leu	Tyr	Asp 80	
CAG	TAC	TGC	AAG	GAC	CAC	TTC	AGC	GAC	GGG	CAC	TGC	GAC	CAG	GGC	TGC	288
Gln	Tyr	Cys	Lys	Asp 85	His	Phe	Ser	Asp	Gly 90	His	Cys	Asp	Gln	Gly 95	Cys	
AAC	AGC	GCG	GAG	TGC	GAG	TGG	GAC	GGG	CTG	GAC	TGT	GCG	GAG	CAT	GTA	336
Asn	Ser	Ala	Glu 100	Cys	Glu	Trp	Asp	Gly 105	Leu	Asp	Cys	Ala	Glu 110	His	Val	
CCC	GAG	AGG	CTG	GCG	GCC	GGC	ACG	CTG	GTG	GTG	GTG	GTG	CTG	ATG	CCG	384
Pro	Glu	Arg 115	Leu	Ala	Ala	Gly	Thr 120	Leu	Val	Val	Val	Val	Leu	Met	Pro	
CCG	GAG	CAG	CTG	GCG	AAC	AGC	TCC	TTC	CAC	TTC	CTG	CGG	GAG	CTC	AGC	432
Pro	Glu	Gln	Leu	Arg	Asn 130	Ser	Ser 135	Phe	His	Phe	Leu 140	Arg	Glu	Leu	Ser	
CGC	GTG	CTG	CAC	ACC	AAC	GTG	GTC	TTC	AAG	CGT	GAC	GCA	CAC	GGC	CAG	480
Arg	Val	Leu	His	Thr 145	Asn 150	Val	Val	Phe	Lys 155	Arg	Asp	Ala	His	Gly 160	Gln	
CAG	ATG	ATC	TTC	CCC	TAC	TAC	GGC	CGC	GAG	GAG	GAG	CTG	CGC	AAG	CAC	528
Gln	Met	Ile	Phe	Pro 165	Tyr	Tyr	Gly	Arg	Glu 170	Glu	Glu	Leu	Arg	Lys 175	His	
CCC	ATC	AAG	CGT	GCC	GCC	GAG	GGC	TGG	GCC	GCA	CCT	GAC	GCC	CTG	CTG	576
Pro	Ile	Lys	Arg 180	Ala	Ala	Glu	Gly 185	Trp	Ala	Ala	Pro	Asp	Ala 190	Leu	Leu	
GGC	CAG	GTG	AAG	GCC	TCG	CTG	CTC	CCT	GGT	GGC	AGC	GAG	GGT	GGG	CGG	624
Gly	Gln	Val 195	Lys	Ala	Ser	Leu	Leu 200	Pro	Gly	Gly	Ser	Glu 205	Gly	Gly	Arg	
CGG	CGG	AGG	GAG	CTG	GAC	CCC	ATG	GAC	GTC	CGC	GGC	TCC	ATC	GTC	TAC	672
Arg	Arg 210	Arg	Glu	Leu	Asp	Pro 215	Met	Asp	Val	Arg	Gly 220	Ser	Ile	Val	Tyr	
CTG	GAG	ATT	GAC	AAC	CGG	CAG	TGT	GTG	CAG	GCC	TCC	TCG	CAG	TGC	TTC	720
Leu	Glu	Ile	Asp	Asn 225	Arg 230	Gln	Cys	Val	Gln 235	Ala	Ser	Ser	Gln	Cys 240	Phe	
CAG	AGT	GCC	ACC	GAC	GTG	GCC	GCA	TTC	CTG	GGA	GCG	CTC	GCC	TCG	CTG	768
Gln	Ser	Ala	Thr 245	Asp	Val	Ala	Ala	Phe 250	Leu	Gly	Ala	Leu	Ala 255	Ser	Leu	
GGC	AGC	CTC	AAC	ATC	CCC	TAC	AAG	ATC	GAG	GCC	GTG	CAG	AGT	GAG	ACC	816
Gly	Ser	Leu	Asn 260	Ile	Pro	Tyr	Lys 265	Ile	Glu	Ala	Val	Gln 270	Ser	Glu	Thr	
GTG	GAG	CCG	CCC	CCG	CCG	GCG	CAG	CTG	CAC	TTC	ATG	TAC	GTG	GCG	GCG	864
Val	Glu	Pro 275	Pro	Pro	Pro	Ala	Gln 280	Leu	His	Phe	Met	Tyr 285	Val	Ala	Ala	
GCC	GCC	TTT	GTG	CTT	CTG	TTC	TTC	GTG	GGC	TGC	GGG	GTG	CTG	CTG	TCC	912
Ala	Ala	Phe 290	Val	Leu	Leu 295	Phe	Phe	Val	Gly	Cys 300	Gly	Val	Leu	Leu	Ser	
CGC	AAG	CGC	CGG	CGG	CAG	CAT	GGC	CAG	CTC	TGG	TTC	CCT	GAG	GGC	TTC	960
Arg	Lys	Arg	Arg	Arg 305	Gln 310	His	Gly	Gln	Leu 315	Trp	Phe	Pro	Glu	Gly 320	Phe	
AAA	GTG	TCT	GAG	GCC	AGC	AAG	AAG	AAG	CGG	CGG	GAG	CCC	CTC	GGC	GAG	1008
Lys	Val	Ser	Glu 325	Ala	Ser	Lys	Lys	Lys	Arg 330	Arg	Glu	Pro	Leu	Gly 335	Glu	

-102-

GAC Asp	TCC Ser	GTG Val	GGC Gly 340	CTC Leu	AAG Lys	CCC Pro	CTG Leu	AAG Lys 345	AAC Asn	GCT Ala	TCA Ser	GAC Asp	GGT Gly 350	GCC Ala	CTC Leu	1056
ATG Met	GAC Asp	GAC Asp 355	AAC Asn	CAG Gln	AAT Asn	GAG Glu	TGG Trp 360	GGG Gly	GAC Asp	GAG Glu	GAC Asp	CTG Leu 365	GAG Glu	ACC Thr	AAG Lys	1104
AAG Lys	TTC Phe 370	CGG Arg	TTC Phe	GAG Glu	GAG Glu	CCC Pro 375	GTG Val	GTT Val	CTG Leu	CCT Pro	GAC Asp 380	CTG Leu	GAC Asp	GAC Asp	CAG Gln	1152
ACA Thr 385	GAC Asp	CAC His	CGG Arg	CAG Gln	TGG Trp 390	ACT Thr	CAG Gln	CAG Gln	CAC His	CTG Leu 395	GAT Asp	GCC Ala	GCT Ala	GAC Asp	CTG Leu 400	1200
CGC Arg	ATG Met	TCT Ser	GCC Ala 405	ATG Met	GCC Ala	CCC Pro	ACA Thr	CCG Pro	CCC Pro	CAG Gln	GGT Gly	GAG Glu	GTT Val	GAC Asp 415	GCC Ala	1248
GAC Asp	TGC Cys	ATG Met	GAC Asp 420	GTC Val	AAT Asn	GTC Val	CGC Arg	GGG Gly 425	CCT Pro	GAT Asp	GGC Gly	TTC Phe	ACC Thr	CCG Pro	CTC Leu	1296
ATG Met	ATC Ile	GCC Ala 435	TCC Ser	TGC Cys	AGC Ser	GGG Gly	GGC Gly 440	GGC Gly	CTG Leu	GAG Glu	ACG Thr	GGC Gly 445	AAC Asn	AGC Ser	GAG Glu	1344
GAA Glu 450	GAG Glu	GAG Glu	GAC Asp	GCG Ala	CCG Pro	GCC Ala 455	GTC Val	ATC Ile	TCC Ser	GAC Asp	TTC Phe	ATC Ile	TAC Tyr	CAG Gln	GGC Gly	1392
GCC Ala 465	AGC Ser	CTG Leu	CAC His	AAC Asn	CAG Gln 470	ACA Thr	GAC Asp	CGC Arg	ACG Thr	GGC Gly 475	GAG Glu	ACC Thr	GCC Ala	TTG Leu	CAC His 480	1440
CTG Leu	GCC Ala	GCC Ala	CGC Arg 485	TAC Tyr	TCA Ser	CGC Arg	TCT Ser	GAT Asp	GCC Ala 490	GCC Ala	AAG Lys	CGC Arg	CTG Leu	CTG Leu 495	GAG Glu	1488
GCC Ala 500	AGC Ser	GCA Ala	GAT Asp	GCC Ala	AAC Asn	ATC Ile	CAG Gln	GAC Asp 505	AAC Asn	ATG Met	GGC Gly	CGC Arg	ACC Thr	CCG Pro	CTG Leu	1536
CAT His	GCG Ala	GCT Ala 515	GTG Val	TCT Ser	GCC Ala	GAC Asp	GCA Ala 520	CAA Gln	GGT Gly	GTC Val	TTC Phe	CAG Gln 525	ATC Ile	CTG Leu	ATC Ile	1584
CGG Arg 530	AAC Asn	CGA Arg	GCC Ala	ACA Thr	GAC Asp	CTG Leu 535	GAT Asp	GCC Ala	CGC Arg	ATG Met	CAT His 540	GAT Asp	GGC Gly	ACG Thr	ACG Thr	1632
CCA Pro 545	CTG Leu	ATC Ile	CTG Leu	GCT Ala	GCC Ala 550	CGC Arg	CTG Leu	GCC Ala	GTG Val	GAG Glu 555	GGC Gly	ATG Met	CTG Leu	GAG Glu	GAC Asp 560	1680
CTC Leu	ATC Ile	AAC Asn	TCA Ser	CAC His 565	GCC Ala	GAC Asp	GTC Val	AAC Asn	GCC Ala 570	GTA Val	GAT Asp	GAC Asp	CTG Leu	GGC Gly 575	AAG Lys	1728
TCC Ser	GCC Ala	CTG Leu	CAC His 580	TGG Trp	GCC Ala	GCC Ala	GCC Ala	GTG Val 585	AAC Asn	AAT Asn	GTG Val	GAT Asp	GCC Ala 590	GCA Ala	GTT Val	1776
GTG Val	CTC Leu	CTG Lys 595	AAG Lys	AAC Asn	GGG Gly	GCT Ala	AAC Asn 600	AAA Lys	GAT Asp	ATG Met	CAG Gln	AAC Asn 605	AAC Asn	AGG Arg	GAG Glu	1824

-103-

GAG Glu 610	ACA Thr 610	CCC Pro 610	CTG Leu 610	TTT Phe 610	CTG Leu 615	GCC Ala 615	GCC Ala 615	CGG Arg 615	GAG Glu 620	GGC Gly 620	AGC Ser 620	TAC Tyr 620	GAG Glu 620	ACC Thr 620	GCC Ala 620	1872
AAG Lys 625	GTG Val 625	CTG Leu 625	CTG Leu 625	GAC Asp 630	CAC His 630	TTT Phe 630	GCC Ala 630	AAC Asn 630	CGG Arg 635	GAC Asp 635	ATC Ile 635	ACG Thr 635	GAT Asp 640	CAT His 640	ATG Met 640	1920
GAC Asp 645	CGC Arg 645	CTG Leu 645	CCG Pro 645	CGC Arg 645	GAC Asp 645	ATC Ile 645	GCA Ala 645	CAG Gln 645	GAG Glu 650	CGC Arg 650	ATG Met 650	CAT His 655	CAC His 655	GAC Asp 655	ATC Ile 655	1968
GTG Val 660	AGG Arg 660	CTG Leu 660	CTG Leu 660	GAC Asp 660	GAG Glu 660	TAC Tyr 660	AAC Asn 665	CTG Leu 665	GTG Val 665	CGC Arg 665	AGC Ser 665	CCG Pro 670	CAG Gln 670	CTG Leu 670	CAC His 670	2016
GGA Gly 675	GCC Ala 675	CCG Pro 675	CTG Leu 675	GGG Gly 675	GGC Gly 675	ACG Thr 675	CCC Pro 680	ACC Thr 680	CTG Leu 680	TCG Ser 680	CCC Pro 685	CCG Pro 685	CTC Leu 685	TGC Cys 685	TCG Ser 685	2064
CCC Pro 690	AAC Asn 690	GGC Gly 690	TAC Tyr 690	CTG Leu 690	GGC Gly 695	AGC Ser 695	CTC Leu 695	AAG Lys 695	CCC Pro 700	GGC Gly 700	GTG Val 700	CAG Gln 700	GGC Gly 700	AAG Lys 700	AAG Lys 700	2112
GTC Val 705	CGC Arg 705	AAG Lys 705	CCC Pro 705	AGC Ser 710	AGC Ser 710	AAA Lys 710	GGC Gly 710	CTG Leu 710	GCC Ala 715	TGT Cys 715	GGA Gly 715	AGC Ser 715	AAG Lys 715	GAG Glu 720	GCC Ala 720	2160
AAG Lys 725	GAC Asp 725	CTC Leu 725	AAG Lys 725	GCA Ala 725	CGG Arg 725	AGG Arg 725	AAG Lys 730	AAG Lys 730	TCC Ser 730	CAG Gln 730	GAT Asp 730	GGC Gly 735	AAG Lys 735	GGC Gly 735	TGC Cys 735	2208
CTG Leu 740	CTG Leu 740	GAC Asp 740	AGC Ser 740	TCC Ser 740	GGC Gly 745	ATG Met 745	CTC Leu 745	TCG Ser 745	CCC Pro 745	GTG Val 745	GAC Asp 745	TCC Ser 750	CTG Leu 750	GAG Glu 750	TCA Ser 750	2256
CCC Pro 755	CAT His 755	GGC Gly 755	TAC Tyr 755	CTG Leu 755	TCA Ser 760	GAC Asp 760	GTG Val 760	GCC Ala 760	TCG Ser 765	CCG Pro 765	CCA Pro 765	CTG Leu 765	CTG Leu 765	CCC Pro 765	TCC Ser 765	2304
CCG Pro 770	TTC Phe 770	CAG Gln 770	CAG Gln 770	TCT Ser 775	CCG Pro 775	TCC Ser 775	GTG Val 775	CCC Pro 775	CTC Leu 780	AAC Asn 780	CAC His 780	CTG Leu 780	CCT Pro 780	GGG Gly 780	ATG Met 780	2352
CCC Pro 785	GAC Asp 785	ACC Thr 785	CAC His 785	CTG Leu 790	GGC Gly 790	ATC Ile 790	GGG Gly 795	CAC His 795	CTG Leu 795	AAC Val 795	GTG Val 795	GCG Ala 795	GCC Ala 795	AAG Lys 800	CCC Pro 800	2400
GAG Glu 805	ATG Met 805	GCG Ala 805	GCG Ala 805	CTG Leu 805	GGT Gly 805	GGG Gly 805	GGC Gly 810	GGC Gly 810	CGG Arg 810	CTG Leu 810	GCC Ala 810	TTT Phe 815	GAG Glu 815	ACT Thr 815	GGC Gly 815	2448
CCA Pro 820	CCT Pro 820	CGT Arg 820	CTC Leu 820	TCC Ser 820	CAC His 820	CTG Leu 825	CCT Pro 825	GTG Val 825	GCC Ala 825	TCT Ser 830	GGC Gly 830	ACC Thr 830	AGC Ser 830	ACC Thr 830	GTC Val 830	2496
CTG Leu 835	GGC Gly 835	TCC Ser 835	AGC Ser 835	AGC Ser 835	GGA Gly 840	GGG Gly 840	GCC Ala 840	CTG Leu 840	AAT Asn 845	TTC Phe 845	ACT Thr 845	GTG Val 845	GGC Gly 845	GGG Gly 845	TCC Ser 845	2544
ACC Thr 850	AGT Ser 850	TTG Leu 850	AAT Asn 850	GGT Gly 855	CAA Gln 855	TGC Cys 855	GAG Glu 855	TGG Trp 855	CTG Leu 860	TCC Ser 860	CGG Arg 860	CTG Leu 860	CAG Gln 860	AGC Ser 860	GGC Gly 860	2592
ATG Met 865	GTG Val 865	CCG Pro 865	AAC Asn 865	CAA Gln 870	TAC Tyr 870	AAC Asn 870	CCT Pro 875	CTG Leu 875	CGG Arg 875	GGG Gly 875	AGT Ser 875	GTG Val 875	GCA Ala 875	CCA Pro 880	GGC Gly 880	2640

-104-

CCC CTG AGC ACA CAG GCC CCC TCC CTG CAG CAT GGC ATG GTA GGC CCG Pro Leu Ser Thr Gln Ala Pro Ser Leu Gln His Gly Met Val Gly Pro 885 890 895	2688
CTG CAC AGT AGC CTT GCT GCC AGC GCC CTG TCC CAG ATG ATG AGC TAC Leu His Ser Ser Leu Ala Ala Ser Ala Leu Ser Gln Met Met Ser Tyr 900 905 910	2736
CAG GGC CTG CCC AGC ACC CGG CTG GCC ACC CAG CCT CAC CTG GTG CAG Gln Gly Leu Pro Ser Thr Arg Leu Ala Thr Gln Pro His Leu Val Gln 915 920 925	2784
ACC CAG CAG GTG CAG CCA CAA AAC TTA CAG ATG CAG CAG CAG AAC CTG Thr Gln Gln Val Gln Pro Gln Asn Leu Gln Met Gln Gln Gln Asn Leu 930 935 940	2832
CAG CCA GCA AAC ATC CAG CAG CAG CAA AGC CTG CAG CCG CCA CCA CCA Gln Pro Ala Asn Ile Gln Gln Gln Gln Ser Leu Gln Pro Pro Pro Pro 945 950 955 960	2880
CCA CCA CAG CCG CAC CTT GGC GTG AGC TCA GCA GCC AGC GGC CAC CTG Pro Pro Gln Pro His Leu Gly Val Ser Ser Ala Ala Ser Gly His Leu 965 970 975	2928
GGC CGG AGC TTC CTG AGT GGA GAG CCG AGC CAG GCA GAC GTG CAG CCA Gly Arg Ser Phe Leu Ser Gly Glu Pro Ser Gln Ala Asp Val Gln Pro 980 985 990	2976
CTG GGC CCC AGC AGC CTG GCG GTG CAC ACT ATT CTG CCC CAG GAG AGC Leu Gly Pro Ser Ser Leu Ala Val His Thr Ile Leu Pro Gln Glu Ser 995 1000 1005	3024
CCC GCC CTG CCC ACG TCG CTG CCA TCC TCG CTG GTC CCA CCC GTG ACC Pro Ala Leu Pro Thr Ser Leu Pro Ser Ser Leu Val Pro Pro Val Thr 1010 1015 1020	3072
GCA GCC CAG TTC CTG ACG CCC CCC TCG CAG CAC AGC TAC TCC TCG CCT Ala Ala Gln Phe Leu Thr Pro Pro Ser Gln His Ser Tyr Ser Ser Pro 1025 1030 1035 1040	3120
GTG GAC AAC ACC CCC AGC CAC CAG CTA CAG GTG CCT GTT CCT GTA ATG Val Asp Asn Thr Pro Ser His Gln Leu Gln Val Pro Val Pro Val Met 1045 1050 1055	3168
GTA ATG ATC CGA TCT TCG GAT CCT TCT AAA GGC TCA TCA ATT TTG ATC Val Met Ile Arg Ser Ser Asp Pro Ser Lys Gly Ser Ser Ile Leu Ile 1060 1065 1070	3216
GAA GCT CCC GAC TCA TGG Glu Ala Pro Asp Ser Trp 1075	3234

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1078 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Cys Gln Glu Asp Ala Gly Asn Lys Val Cys Ser Leu Gln Cys Asn Asn  
1 5 10 15

-105-

His Ala Cys Gly Trp Asp Gly Gly Asp Cys Ser Leu Asn Phe Asn Asp  
 20 25 30  
 Pro Trp Lys Asn Cys Thr Gln Ser Leu Gln Cys Trp Lys Tyr Phe Ser  
 35 40 45  
 Asp Gly His Cys Asp Ser Gln Cys Asn Ser Ala Gly Cys Leu Phe Asp  
 50 55 60  
 Gly Phe Asp Cys Gln Arg Ala Glu Gly Gln Cys Asn Pro Leu Tyr Asp  
 65 70 75 80  
 Gln Tyr Cys Lys Asp His Phe Ser Asp Gly His Cys Asp Gln Gly Cys  
 85 90 95  
 Asn Ser Ala Glu Cys Glu Trp Asp Gly Leu Asp Cys Ala Glu His Val  
 100 105 110  
 Pro Glu Arg Leu Ala Ala Gly Thr Leu Val Val Val Val Leu Met Pro  
 115 120 125  
 Pro Glu Gln Leu Arg Asn Ser Ser Phe His Phe Leu Arg Glu Leu Ser  
 130 135 140  
 Arg Val Leu His Thr Asn Val Val Phe Lys Arg Asp Ala His Gly Gln  
 145 150 155 160  
 Gln Met Ile Phe Pro Tyr Tyr Gly Arg Glu Glu Glu Leu Arg Lys His  
 165 170 175  
 Pro Ile Lys Arg Ala Ala Glu Gly Trp Ala Ala Pro Asp Ala Leu Leu  
 180 185 190  
 Gly Gln Val Lys Ala Ser Leu Leu Pro Gly Gly Ser Glu Gly Gly Arg  
 195 200 205  
 Arg Arg Arg Glu Leu Asp Pro Met Asp Val Arg Gly Ser Ile Val Tyr  
 210 215 220  
 Leu Glu Ile Asp Asn Arg Gln Cys Val Gln Ala Ser Ser Gln Cys Phe  
 225 230 235 240  
 Gln Ser Ala Thr Asp Val Ala Ala Phe Leu Gly Ala Leu Ala Ser Leu  
 245 250 255  
 Gly Ser Leu Asn Ile Pro Tyr Lys Ile Glu Ala Val Gln Ser Glu Thr  
 260 265 270  
 Val Glu Pro Pro Pro Pro Ala Gln Leu His Phe Met Tyr Val Ala Ala  
 275 280 285  
 Ala Ala Phe Val Leu Leu Phe Phe Val Gly Cys Gly Val Leu Leu Ser  
 290 295 300  
 Arg Lys Arg Arg Arg Gln His Gly Gln Leu Trp Phe Pro Glu Gly Phe  
 305 310 315 320  
 Lys Val Ser Glu Ala Ser Lys Lys Lys Arg Arg Glu Pro Leu Gly Glu  
 325 330 335  
 Asp Ser Val Gly Leu Lys Pro Leu Lys Asn Ala Ser Asp Gly Ala Leu  
 340 345 350  
 Met Asp Asp Asn Gln Asn Glu Trp Gly Asp Glu Asp Leu Glu Thr Lys  
 355 360 365  
 Lys Phe Arg Phe Glu Glu Pro Val Val Leu Pro Asp Leu Asp Asp Gln

-106-

370					375					380					
Thr	Asp	His	Arg	Gln	Trp	Thr	Gln	Gln	His	Leu	Asp	Ala	Ala	Asp	Leu
385					390					395					400
Arg	Met	Ser	Ala	Met	Ala	Pro	Thr	Pro	Pro	Gln	Gly	Glu	Val	Asp	Ala
				405					410					415	
Asp	Cys	Met	Asp	Val	Asn	Val	Arg	Gly	Pro	Asp	Gly	Phe	Thr	Pro	Leu
			420					425					430		
Met	Ile	Ala	Ser	Cys	Ser	Gly	Gly	Gly	Leu	Glu	Thr	Gly	Asn	Ser	Glu
		435					440						445		
Glu	Glu	Glu	Asp	Ala	Pro	Ala	Val	Ile	Ser	Asp	Phe	Ile	Tyr	Gln	Gly
	450					455					460				
Ala	Ser	Leu	His	Asn	Gln	Thr	Asp	Arg	Thr	Gly	Glu	Thr	Ala	Leu	His
465					470					475					480
Leu	Ala	Ala	Arg	Tyr	Ser	Arg	Ser	Asp	Ala	Ala	Lys	Arg	Leu	Leu	Glu
				485					490					495	
Ala	Ser	Ala	Asp	Ala	Asn	Ile	Gln	Asp	Asn	Met	Gly	Arg	Thr	Pro	Leu
			500					505					510		
His	Ala	Ala	Val	Ser	Ala	Asp	Ala	Gln	Gly	Val	Phe	Gln	Ile	Leu	Ile
			515				520					525			
Arg	Asn	Arg	Ala	Thr	Asp	Leu	Asp	Ala	Arg	Met	His	Asp	Gly	Thr	Thr
	530					535					540				
Pro	Leu	Ile	Leu	Ala	Ala	Arg	Leu	Ala	Val	Glu	Gly	Met	Leu	Glu	Asp
545					550					555					560
Leu	Ile	Asn	Ser	His	Ala	Asp	Val	Asn	Ala	Val	Asp	Asp	Leu	Gly	Lys
				565				570						575	
Ser	Ala	Leu	His	Trp	Ala	Ala	Ala	Val	Asn	Asn	Val	Asp	Ala	Ala	Val
			580					585					590		
Val	Leu	Leu	Lys	Asn	Gly	Ala	Asn	Lys	Asp	Met	Gln	Asn	Asn	Arg	Glu
	595						600					605			
Glu	Thr	Pro	Leu	Phe	Leu	Ala	Ala	Arg	Glu	Gly	Ser	Tyr	Glu	Thr	Ala
	610					615					620				
Lys	Val	Leu	Leu	Asp	His	Phe	Ala	Asn	Arg	Asp	Ile	Thr	Asp	His	Met
625					630					635					640
Asp	Arg	Leu	Pro	Arg	Asp	Ile	Ala	Gln	Glu	Arg	Met	His	His	Asp	Ile
				645					650					655	
Val	Arg	Leu	Leu	Asp	Glu	Tyr	Asn	Leu	Val	Arg	Ser	Pro	Gln	Leu	His
			660					665					670		
Gly	Ala	Pro	Leu	Gly	Gly	Thr	Pro	Thr	Leu	Ser	Pro	Pro	Leu	Cys	Ser
		675					680					685			
Pro	Asn	Gly	Tyr	Leu	Gly	Ser	Leu	Lys	Pro	Gly	Val	Gln	Gly	Lys	Lys
	690					695					700				
Val	Arg	Lys	Pro	Ser	Ser	Lys	Gly	Leu	Ala	Cys	Gly	Ser	Lys	Glu	Ala
705					710					715					720
Lys	Asp	Leu	Lys	Ala	Arg	Arg	Lys	Lys	Ser	Gln	Asp	Gly	Lys	Gly	Cys
			725					730						735	



-107-

Leu Leu Asp Ser Ser Gly Met Leu Ser Pro Val Asp Ser Leu Glu Ser  
 740 745 750  
 Pro His Gly Tyr Leu Ser Asp Val Ala Ser Pro Pro Leu Leu Pro Ser  
 755 760 765  
 Pro Phe Gln Gln Ser Pro Ser Val Pro Leu Asn His Leu Pro Gly Met  
 770 775 780  
 Pro Asp Thr His Leu Gly Ile Gly His Leu Asn Val Ala Ala Lys Pro  
 785 790 795 800  
 Glu Met Ala Ala Leu Gly Gly Gly Gly Arg Leu Ala Phe Glu Thr Gly  
 805 810 815  
 Pro Pro Arg Leu Ser His Leu Pro Val Ala Ser Gly Thr Ser Thr Val  
 820 825 830  
 Leu Gly Ser Ser Ser Gly Gly Ala Leu Asn Phe Thr Val Gly Gly Ser  
 835 840 845  
 Thr Ser Leu Asn Gly Gln Cys Glu Trp Leu Ser Arg Leu Gln Ser Gly  
 850 855 860  
 Met Val Pro Asn Gln Tyr Asn Pro Leu Arg Gly Ser Val Ala Pro Gly  
 865 870 875 880  
 Pro Leu Ser Thr Gln Ala Pro Ser Leu Gln His Gly Met Val Gly Pro  
 885 890 895  
 Leu His Ser Ser Leu Ala Ala Ser Ala Leu Ser Gln Met Met Ser Tyr  
 900 905 910  
 Gln Gly Leu Pro Ser Thr Arg Leu Ala Thr Gln Pro His Leu Val Gln  
 915 920 925  
 Thr Gln Gln Val Gln Pro Gln Asn Leu Gln Met Gln Gln Gln Asn Leu  
 930 935 940  
 Gln Pro Ala Asn Ile Gln Gln Gln Gln Ser Leu Gln Pro Pro Pro Pro  
 945 950 955 960  
 Pro Pro Gln Pro His Leu Gly Val Ser Ser Ala Ala Ser Gly His Leu  
 965 970 975  
 Gly Arg Ser Phe Leu Ser Gly Glu Pro Ser Gln Ala Asp Val Gln Pro  
 980 985 990  
 Leu Gly Pro Ser Ser Leu Ala Val His Thr Ile Leu Pro Gln Glu Ser  
 995 1000 1005  
 Pro Ala Leu Pro Thr Ser Leu Pro Ser Ser Leu Val Pro Pro Val Thr  
 1010 1015 1020  
 Ala Ala Gln Phe Leu Thr Pro Pro Ser Gln His Ser Tyr Ser Ser Pro  
 1025 1030 1035 1040  
 Val Asp Asn Thr Pro Ser His Gln Leu Gln Val Pro Val Pro Val Met  
 1045 1050 1055  
 Val Met Ile Arg Ser Ser Asp Pro Ser Lys Gly Ser Ser Ile Leu Ile  
 1060 1065 1070  
 Glu Ala Pro Asp Ser Trp  
 1075

(2) INFORMATION FOR SEQ ID NO:12:

-108-

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4268 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 2..1972

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

G GAG GTG GAT GTG TTA GAT GTG AAT GTC CGT GGC CCA GAT GGC TGC	46
Glu Val Asp Val Leu Asp Val Asn Val Arg Gly Pro Asp Gly Cys	
1 5 10 15	
ACC CCA TTG ATG TTG GCT TCT CTC CGA GGA GGC AGC TCA GAT TTG AGT	94
Thr Pro Leu Met Leu Ala Ser Leu Arg Gly Gly Ser Ser Asp Leu Ser	
20 25 30	
GAT GAA GAT GAA GAT GCA GAG GAC TCT TCT GCT AAC ATC ATC ACA GAC	142
Asp Glu Asp Glu Asp Ala Glu Asp Ser Ser Ala Asn Ile Ile Thr Asp	
35 40 45	
TTG GTC TAC CAG GGT GCC AGC CTC CAG GCC CAG ACA GAC CGG ACT GGT	190
Leu Val Tyr Gln Gly Ala Ser Leu Gln Ala Gln Thr Asp Arg Thr Gly	
50 55 60	
GAG ATG GCC CTG CAC CTT GCA GCC CGC TAC TCA CGG GCT GAT GCT GCC	238
Glu Met Ala Leu His Leu Ala Ala Arg Tyr Ser Arg Ala Asp Ala Ala	
65 70 75	
AAG CGT CTC CTG GAT GCA GGT GCA GAT GCC AAT GCC CAG GAC AAC ATG	286
Lys Arg Leu Leu Asp Ala Gly Ala Asp Ala Asn Ala Gln Asp Asn Met	
80 85 90 95	
GGC CGC TGT CCA CTC CAT GCT GCA GTG GCA GCT GAT GCC CAA GGT GTC	334
Gly Arg Cys Pro Leu His Ala Ala Val Ala Ala Asp Ala Gln Gly Val	
100 105 110	
TTC CAG ATT CTG ATT CGC AAC CGA GTA ACT GAT CTA GAT GCC AGG ATG	382
Phe Gln Ile Leu Ile Arg Asn Arg Val Thr Asp Leu Asp Ala Arg Met	
115 120 125	
AAT GAT GGT ACT ACA CCC CTG ATC CTG GCT GCC CGC CTG GCT GTG GAG	430
Asn Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala Arg Leu Ala Val Glu	
130 135 140	
GGA ATG GTG GCA GAA CTG ATC AAC TGC CAA GCG GAT GTG AAT GCA GTG	478
Gly Met Val Ala Glu Leu Ile Asn Cys Gln Ala Asp Val Asn Ala Val	
145 150 155	
GAT GAC CAT GGA AAA TCT GCT CTT CAC TGG GCA GCT GCT GTC AAT AAT	526
Asp Asp His Gly Lys Ser Ala Leu His Trp Ala Ala Ala Val Asn Asn	
160 165 170 175	
GTG GAG GCA ACT CTT TTG TTG TTG AAA AAT GGG GCC AAC CGA GAC ATG	574
Val Glu Ala Thr Leu Leu L u Leu Lys Asn Gly Ala Asn Arg Asp Met	
180 185 190	
CAG GAC AAC AAG GAA GAG ACA CCT CTG TTT CTT GCT GCC CGG GAG GGG	622
Gln Asp Asn Lys Glu Glu Thr Pro Leu Phe Leu Ala Ala Arg Glu Gly	
195 200 205	

-109-

AGC TAT GAA GCA GCC AAG ATC CTG TTA GAC CAT TTT GCC AAT CGA GAC Ser Tyr Glu Ala Ala Lys Ile Leu Leu Asp His Phe Ala Asn Arg Asp 210 215 220	670
ATC ACA GAC CAT ATG GAT CGT CTT CCC CGG GAT GTG GCT CGG GAT CGC Ile Thr Asp His Met Asp Arg Leu Pro Arg Asp Val Ala Arg Asp Arg 225 230 235	718
ATG CAC CAT GAC ATT GTG CGC CTT CTG GAT GAA TAC AAT GTG ACC CCA Met His His Asp Ile Val Arg Leu Leu Asp Glu Tyr Asn Val Thr Pro 240 245 250 255	766
AGC CCT CCA GGC ACC GTG TTG ACT TCT GCT CTC TCA CCT GTC ATC TGT Ser Pro Pro Gly Thr Val Leu Thr Ser Ala Leu Ser Pro Val Ile Cys 260 265 270	814
GGG CCC AAC AGA TCT TTC CTC AGC CTG AAG CAC ACC CCA ATG GGC AAG Gly Pro Asn Arg Ser Phe Leu Ser Leu Lys His Thr Pro Met Gly Lys 275 280 285	862
AAG TCT AGA CGG CCC AGT GCC AAG AGT ACC ATG CCT ACT AGC CTC CCT Lys Ser Arg Arg Pro Ser Ala Lys Ser Thr Met Pro Thr Ser Leu Pro 290 295 300	910
AAC CTT GCC AAG GAG GCA AAG GAT GCC AAG GGT AGT AGG AGG AAG AAG Asn Leu Ala Lys Glu Ala Lys Asp Ala Lys Gly Ser Arg Arg Lys Lys 305 310 315	958
TCT CTG AGT GAG AAG GTC CAA CTG TCT GAG AGT TCA GTA ACT TTA TCC Ser Leu Ser Glu Lys Val Gln Leu Ser Glu Ser Ser Val Thr Leu Ser 320 325 330 335	1006
CCT GTT GAT TCC CTA GAA TCT CCT CAC ACG TAT GTT TCC GAC ACC ACA Pro Val Asp Ser Leu Glu Ser Pro His Thr Tyr Val Ser Asp Thr Thr 340 345 350	1054
TCC TCT CCA ATG ATT ACA TCC CCT GGG ATC TTA CAG GCC TCA CCC AAC Ser Ser Pro Met Ile Thr Ser Pro Gly Ile Leu Gln Ala Ser Pro Asn 355 360 365	1102
CCT ATG TTG GCC ACT GCC GCC CCT CCT GCC CCA GTC CAT GCC CAG CAT Pro Met Leu Ala Thr Ala Ala Pro Pro Ala Pro Val His Ala Gln His 370 375 380	1150
GCA CTA TCT TTT TCT AAC CTT CAT GAA ATG CAG CCT TTG GCA CAT GGG Ala Leu Ser Phe Ser Asn Leu His Glu Met Gln Pro Leu Ala His Gly 385 390 395	1198
GCC AGC ACT GTG CTT CCC TCA GTG AGC CAG TTG CTA TCC CAC CAC CAC Ala Ser Thr Val Leu Pro Ser Val Ser Gln Leu Leu Ser His His His 400 405 410 415	1246
ATT GTG TCT CCA GGC AGT GGC AGT GCT GGA AGC TTG AGT AGG CTC CAT Ile Val Ser Pro Gly Ser Gly Ser Ala Gly Ser Leu Ser Arg Leu His 420 425 430	1294
CCA GTC CCA GTC CCA GCA GAT TGG ATG AAC CGC ATG GAG GTG AAT GAG Pro Val Pro Val Pro Ala Asp Trp Met Asn Arg Met Glu Val Asn Glu 435 440 445	1342
ACC CAG TAC AAT GAG ATG TTT GGT ATG GTC CTG GCT CCA GCT GAG GGC Thr Gln Tyr Asn Glu Met Phe Gly Met Val Leu Ala Pro Ala Glu Gly 450 455 460	1390
ACC CAT CCT GCC ATA GCT CCC CAG AGC AGG CCA CCT GAA GGG AAG CAC Thr His Pro Gly Ile Ala Pro Gln Ser Arg Pro Pro Glu Gly Lys His 465 470 475	1438

-110-

ATA ACC ACC CCT CGG GAG CCC TTG CCC CCC ATT GTG ACT TTC CAG CTC Ile Thr Thr Pro Arg Glu Pro Leu Pro Pro Ile Val Thr Phe Gln Leu 480 485 490 495	1486
ATC CCT AAA GGC AGT ATT GCC CAA CCA GCG GGG GCT CCC CAG CCT CAG Ile Pro Lys Gly Ser Ile Ala Gln Pro Ala Gly Ala Pro Gln Pro Gln 500 505 510	1534
TCC ACC TGC CCT CCA GCT GTT GCG GGC CCC CTG CCC ACC ATG TAC CAG Ser Thr Cys Pro Pro Ala Val Ala Gly Pro Leu Pro Thr Met Tyr Gln 515 520 525	1582
ATT CCA GAA ATG GCC CGT TTG CCC AGT GTG GCT TTC CCC ACT GCC ATG Ile Pro Glu Met Ala Arg Leu Pro Ser Val Ala Phe Pro Thr Ala Met 530 535 540	1630
ATG CCC CAG CAG GAC GGG CAG GTA GCT CAG ACC ATT CTC CCA GCC TAT Met Pro Gln Gln Asp Gly Gln Val Ala Gln Thr Ile Leu Pro Ala Tyr 545 550 555	1678
CAT CCT TTC CCA GCC TCT GTG GGC AAG TAC CCC ACA CCC CCT TCA CAG His Pro Phe Pro Ala Ser Val Gly Lys Tyr Pro Thr Pro Pro Ser Gln 560 565 570 575	1726
CAC AGT TAT GCT TCC TCA AAT GCT GCT GAG CGA ACA CCC AGT CAC AGT His Ser Tyr Ala Ser Ser Asn Ala Ala Glu Arg Thr Pro Ser His Ser 580 585 590	1774
GGT CAC CTC CAG GGT GAG CAT CCC TAC CTG ACA CCA TCC CCA GAG TCT Gly His Leu Gln Gly Glu His Pro Tyr Leu Thr Pro Ser Pro Glu Ser 595 600 605	1822
CCT GAC CAG TGG TCA AGT TCA TCA CCC CAC TCT GCT TCT GAC TGG TCA Pro Asp Gln Trp Ser Ser Ser Pro His Ser Ala Ser Asp Trp Ser 610 615 620	1870
GAT GTG ACC ACC AGC CCT ACC CCT GGG GGT GCT GGA GGA GGT CAG CGG Asp Val Thr Thr Ser Pro Thr Pro Gly Gly Ala Gly Gly Gly Gln Arg 625 630 635	1918
GGA CCT GGG ACA CAC ATG TCT GAG CCA CCA CAC AAC AAC ATG CAG GTT Gly Pro Gly Thr His Met Ser Glu Pro Pro His Asn Asn Met Gln Val 640 645 650 655	1966
TAT GCG TGAGAGAGTC CACCTCCAGT GTAGAGACAT AACTGACTTT TGTAATGCT Tyr Ala	2022
GCTGAGGAAC AAATGAAGGT CATCCGGGAG AGAAATGAAG AAATCTCTGG AGCCAGCTTC	2082
TAGAGGTAGG AAAGAGAAGA TGTTCCTTATT CAGATAATGC AAGAGAAGCA ATTCGTCAGT	2142
TTCAGTGGGT ATCTGCAAGG CTTATTGATT ATTCTAATCT AATAAGACAA GTTTGTGGAA	2202
ATGCAAGATG AATACAAGCC TTGGGTCCAT GTTACTCTC TTCTATTTGG AGAATAAGAT	2262
GGATGCTTAT TGAAGCCCAG ACATTCTTGC AGCTTGGACT GCATTTTAAAG CCCTGCAGGC	2322
TTCTGCCATA TCCATGAGAA GATTCTACAC TAGCGTCCTG TTGGGAATTA TGCCCTGGAA	2382
TTCTGCCTGA ATTGACCTAC GCATCTCCTC CTCCTTGGAC ATTCTTTTGT CTTTATTGG	2442
TGCTTTTGGT TTTGCACCTC TCCGTGATTG TAGCCCTACC AGCATGTTAT AGGGCAAGAC	2502
CTTTGTGCTT TTGATCATTC TGGCCCATGA AAGCAACTTT GGTCTCCTTT CCCCTCCTGT	2562
CTTCCCGGTA TCCCTTGGAG TCTCACAAGG TTTACTTTGG TATGTTTCTC AGCACAAACC	2622

-111-

TTTCAAGTAT	GTTGTTTCTT	TGGAAAATGG	ACATACTGTA	TTGTGTTCTC	CTGCATATAT	2682
CATTCCTGGA	GAGAGAAGGG	GAGAAGAATA	CTTTTCTTCA	ACAAATTTTG	GGGGCAGGAG	2742
ATCCCTTCAA	GAGGCTGCAC	CTTAATTTTT	CTTGTCTGTG	TGCAGGTCTT	CATATAAACT	2802
TTACCAGGAA	GAAGGGTGTG	AGTTTGTGTG	TTTTCTGTGT	ATGGGCCTGG	TCAGTGTAAG	2862
GTTTTATCCT	TGATAGTCTA	GTTACTATGA	CCCTCCCCAC	TTTTTTAAAA	CCAGAAAAAG	2922
GTTTGGAATG	TTGGAATGAC	CAAGAGACAA	GTTAACTCGT	GCAAGAGCCA	GTTACCCACC	2982
CACAGGTCCC	CCTACTTCCT	GCCAAGCATT	CCATTGACTG	CCTGTATGGA	ACACATTGTG	3042
CCCAGATCTG	AGCATTCTAG	GCCTGTTTCA	CTCACTCACC	CAGCATATGA	AACTAGTCTT	3102
AACTGTTGAG	CCTTTCCTTT	CATATCCACA	GAAGACACTG	TCTCAAATGT	TGTACCCCTG	3162
CCATTTAGGA	CTGAACTTTC	CTTAGCCCAA	GGGACCCAGT	GACAGTTGTC	TTCCGTTTGT	3222
CAGATGATCA	GTCTCTACTG	ATTATCTTGC	TGCTTAAAGG	CCTGCTCACC	AATCTTTCTT	3282
TCACACCGTG	TGGTCCGTGT	TACTGGTATA	CCCAGTATGT	TCTCACTGAA	GACATGGACT	3342
TTATATGTTT	AAGTGCAGGA	ATTGGAAAGT	TGGACTTGTT	TTCTATGATC	CAAAACAGCC	3402
CTATAAGAAG	GTTGGAAAAG	GAGGAAGTAT	ATAGCAGCCT	TTGCTATTTT	CTGCTACCAT	3462
TTCTTTTCCT	CTGAAGCGGC	CATGACATTC	CCTTTGGCAA	CTAACGTAGA	AACTCAACAG	3522
AACATTTTCC	TTTCCTAGAG	TCACCTTTTA	GATGATAATG	GACAACTATA	GACTTGCTCA	3582
TTGTTTCTAG	TGATTGCCCC	TCACCTGAAT	CCACTCTCTG	TATTCATGCT	CTTGGCAATT	3642
TCTTTGACTT	TCTTTTAAGG	GCAGAAGCAT	TTTAGTTAAT	TGTAGATAAA	GAATAGTTTT	3702
CTTCCTCTTC	TCCTTGGGCC	AGTTAATAAT	TGGTCCATGG	CTCACTGCA	ACTTCCGTCC	3762
AGTGCTGTGA	TGCCCATGAC	ACCTGCAAAA	TAAGTTCTGC	CTGGGCATTT	TGTAGATATT	3822
AACAGGTGAA	TTCCCGACTC	TTTTGGTTTG	AATGACAGTT	CTCATTCCTT	CTATGGCTGC	3882
AAGTATGCAT	CAGTGCTTCC	CACTTACCTG	ATTTGTCTGT	CGGTGGCCCC	ATATGGAAAC	3942
CCTGCGTGTC	TGTTGGCATA	ATAGTTTACA	AATGGTTTTT	TCAGTCCTAT	CCAAATTTAT	4002
TGAACCAACA	AAAATAATTA	CTTCTGCCCT	GAGATAAGCA	GATTAAGTTT	GTTCAATTCT	4062
TGCTTTATTG	TCTCCATGTG	GCAACATTCT	GTCAGCCTCT	TTCATAGTGT	GCAAACATTT	4122
TATCATTCTA	AATGGTGACT	CTCTGCCCTT	GGACCCATTT	ATTATTCACA	GATGGGGAGA	4182
ACCTATCTGC	ATGGACCCTC	ACCATCCTCT	GTGCAGCACA	CACAGTGCAG	GGAGCCAGTG	4242
GCGATGGCGA	TGACTTTCTT	CCCCCTG				4268

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 657 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

-112-

Glu Val Asp Val Leu Asp Val Asn Val Arg Gly Pro Asp Gly Cys Thr  
 1 5 10 15  
 Pro Leu Met Leu Ala Ser Leu Arg Gly Gly Ser Ser Asp Leu Ser Asp  
 20 25 30  
 Glu Asp Glu Asp Ala Glu Asp Ser Ser Ala Asn Ile Ile Thr Asp Leu  
 35 40 45  
 Val Tyr Gln Gly Ala Ser Leu Gln Ala Gln Thr Asp Arg Thr Gly Glu  
 50 55 60  
 Met Ala Leu His Leu Ala Ala Arg Tyr Ser Arg Ala Asp Ala Ala Lys  
 65 70 75 80  
 Arg Leu Leu Asp Ala Gly Ala Asp Ala Asn Ala Gln Asp Asn Met Gly  
 85 90 95  
 Arg Cys Pro Leu His Ala Ala Val Ala Ala Asp Ala Gln Gly Val Phe  
 100 105 110  
 Gln Ile Leu Ile Arg Asn Arg Val Thr Asp Leu Asp Ala Arg Met Asn  
 115 120 125  
 Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala Arg Leu Ala Val Glu Gly  
 130 135 140  
 Met Val Ala Glu Leu Ile Asn Cys Gln Ala Asp Val Asn Ala Val Asp  
 145 150 155 160  
 Asp His Gly Lys Ser Ala Leu His Trp Ala Ala Ala Val Asn Asn Val  
 165 170 175  
 Glu Ala Thr Leu Leu Leu Leu Lys Asn Gly Ala Asn Arg Asp Met Gln  
 180 185 190  
 Asp Asn Lys Glu Glu Thr Pro Leu Phe Leu Ala Ala Arg Glu Gly Ser  
 195 200 205  
 Tyr Glu Ala Ala Lys Ile Leu Leu Asp His Phe Ala Asn Arg Asp Ile  
 210 215 220  
 Thr Asp His Met Asp Arg Leu Pro Arg Asp Val Ala Arg Asp Arg Met  
 225 230 235 240  
 His His Asp Ile Val Arg Leu Leu Asp Glu Tyr Asn Val Thr Pro Ser  
 245 250 255  
 Pro Pro Gly Thr Val Leu Thr Ser Ala Leu Ser Pro Val Ile Cys Gly  
 260 265 270  
 Pro Asn Arg Ser Phe Leu Ser Leu Lys His Thr Pro Met Gly Lys Lys  
 275 280 285  
 Ser Arg Arg Pro Ser Ala Lys Ser Thr Met Pro Thr Ser Leu Pro Asn  
 290 295 300  
 Leu Ala Lys Glu Ala Lys Asp Ala Lys Gly Ser Arg Arg Lys Lys Ser  
 305 310 315 320  
 Leu Ser Glu Lys Val Gln Leu Ser Glu Ser Ser Val Thr Leu Ser Pro  
 325 330 335  
 Val Asp Ser Leu Glu Ser Pro His Thr Tyr Val Ser Asp Thr Thr S r  
 340 345 350  
 Ser Pro Met Ile Thr Ser Pro Gly Ile Leu Gln Ala Ser Pro Asn Pro

-113-

355		360		365											
Met	Leu	Ala	Thr	Ala	Ala	Pro	Pro	Ala	Pro	Val	His	Ala	Gln	His	Ala
370						375					380				
Leu	Ser	Phe	Ser	Asn	Leu	His	Glu	Met	Gln	Pro	Leu	Ala	His	Gly	Ala
385					390					395					400
Ser	Thr	Val	Leu	Pro	Ser	Val	Ser	Gln	Leu	Leu	Ser	His	His	His	Ile
				405					410					415	
Val	Ser	Pro	Gly	Ser	Gly	Ser	Ala	Gly	Ser	Leu	Ser	Arg	Leu	His	Pro
			420					425					430		
Val	Pro	Val	Pro	Ala	Asp	Trp	Met	Asn	Arg	Met	Glu	Val	Asn	Glu	Thr
		435					440					445			
Gln	Tyr	Asn	Glu	Met	Phe	Gly	Met	Val	Leu	Ala	Pro	Ala	Glu	Gly	Thr
450						455					460				
His	Pro	Gly	Ile	Ala	Pro	Gln	Ser	Arg	Pro	Pro	Glu	Gly	Lys	His	Ile
465					470					475					480
Thr	Thr	Pro	Arg	Glu	Pro	Leu	Pro	Pro	Ile	Val	Thr	Phe	Gln	Leu	Ile
				485					490					495	
Pro	Lys	Gly	Ser	Ile	Ala	Gln	Pro	Ala	Gly	Ala	Pro	Gln	Pro	Gln	Ser
			500					505					510		
Thr	Cys	Pro	Pro	Ala	Val	Ala	Gly	Pro	Leu	Pro	Thr	Met	Tyr	Gln	Ile
		515					520					525			
Pro	Glu	Met	Ala	Arg	Leu	Pro	Ser	Val	Ala	Phe	Pro	Thr	Ala	Met	Met
530						535					540				
Pro	Gln	Gln	Asp	Gly	Gln	Val	Ala	Gln	Thr	Ile	Leu	Pro	Ala	Tyr	His
545					550					555					560
Pro	Phe	Pro	Ala	Ser	Val	Gly	Lys	Tyr	Pro	Thr	Pro	Pro	Ser	Gln	His
				565					570					575	
Ser	Tyr	Ala	Ser	Ser	Asn	Ala	Ala	Glu	Arg	Thr	Pro	Ser	His	Ser	Gly
			580					585					590		
His	Leu	Gln	Gly	Glu	His	Pro	Tyr	Leu	Thr	Pro	Ser	Pro	Glu	Ser	Pro
		595					600					605			
Asp	Gln	Trp	Ser	Ser	Ser	Ser	Pro	His	Ser	Ala	Ser	Asp	Trp	Ser	Asp
610						615					620				
Val	Thr	Thr	Ser	Pro	Thr	Pro	Gly	Gly	Ala	Gly	Gly	Gly	Gln	Arg	Gly
625					630					635					640
Pro	Gly	Thr	His	Met	Ser	Glu	Pro	Pro	His	Asn	Asn	Met	Gln	Val	Tyr
				645					650					655	

Ala

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 77 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

-114-

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Glu Asp Ile Asp Glu Cys Asp Gln Gly Ser Pro Cys Glu His Asn Gly  
 1 5 10 15  
 Ile Cys Val Asn Thr Pro Gly Ser Tyr Arg Cys Asn Cys Ser Gln Gly  
 20 25 30  
 Phe Thr Gly Pro Arg Cys Glu Thr Asn Ile Asn Glu Cys Glu Ser His  
 35 40 45  
 Pro Cys Gln Asn Glu Gly Ser Cys Leu Asp Asp Pro Gly Thr Phe Arg  
 50 55 60  
 Cys Val Cys Met Pro Gly Phe Thr Gly Thr Gln Cys Glu  
 65 70 75

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 78 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Asn Asp Val Asp Glu Cys Ser Leu Gly Ala Asn Pro Cys Glu His Gly  
 1 5 10 15  
 Gly Arg Cys Thr Asn Thr Leu Gly Ser Phe Gln Cys Asn Cys Pro Gln  
 20 25 30  
 Gly Tyr Ala Gly Pro Arg Cys Glu Ile Asp Val Asn Glu Cys Leu Ser  
 35 40 45  
 Asn Pro Cys Gln Asn Asp Ser Thr Cys Leu Asp Gln Ile Gly Glu Phe  
 50 55 60  
 Gln Cys Ile Cys Met Pro Gly Tyr Glu Gly Leu Tyr Cys Glu  
 65 70 75

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 654 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Thr Pro Pro Gln Gly Glu Ile Glu Ala Asp Cys Met Asp Val Asn Val



-115-

1	5	10	15
Arg Gly Pro Asp Gly Phe Thr Pro Leu Met Ile Ala Ser Cys Ser Gly	20	25	30
Gly Gly Leu Glu Thr Gly Asn Ser Glu Glu Glu Glu Asp Ala Ser Ala	35	40	45
Asn Met Ile Ser Asp Phe Ile Gly Gln Gly Ala Gln Leu His Asn Gln	50	55	60
Thr Asp Arg Thr Gly Glu Thr Ala Leu His Leu Ala Ala Arg Tyr Ala	65	70	75
Arg Ala Asp Ala Ala Lys Arg Leu Leu Glu Ser Ser Ala Asp Ala Asn	85	90	95
Val Gln Asp Asn Met Gly Arg Thr Pro Leu His Ala Ala Val Ala Ala	100	105	110
Asp Ala Gln Gly Val Phe Gln Ile Leu Ile Arg Asn Arg Ala Thr Asp	115	120	125
Leu Asp Ala Arg Met Phe Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala	130	135	140
Arg Leu Ala Val Glu Gly Met Val Glu Glu Leu Ile Asn Ala His Ala	145	150	155
Asp Val Asn Ala Val Asp Glu Phe Gly Lys Ser Ala Leu His Trp Ala	165	170	175
Ala Ala Val Asn Asn Val Asp Ala Ala Ala Val Leu Leu Lys Asn Ser	180	185	190
Ala Asn Lys Asp Met Gln Asn Asn Lys Glu Glu Thr Ser Leu Phe Leu	195	200	205
Ala Ala Arg Glu Gly Ser Tyr Glu Thr Ala Lys Val Leu Leu Asp His	210	215	220
Tyr Ala Asn Arg Asp Ile Thr Asp His Met Asp Arg Leu Pro Arg Asp	225	230	235
Ile Ala Gln Glu Arg Met His His Asp Ile Val His Leu Leu Asp Glu	245	250	255
Tyr Asn Leu Val Lys Ser Pro Thr Leu His Asn Gly Pro Leu Gly Ala	260	265	270
Thr Thr Leu Ser Pro Pro Ile Cys Ser Pro Asn Gly Tyr Met Gly Asn	275	280	285
Met Lys Pro Ser Val Gln Ser Lys Lys Ala Arg Lys Pro Ser Ile Lys	290	295	300
Gly Asn Gly Cys Lys Glu Ala Lys Glu Leu Lys Ala Arg Arg Lys Lys	305	310	315
Ser Gln Asp Gly Lys Thr Thr Leu Leu Asp Ser Gly Ser Ser Gly Val	325	330	335
Leu Ser Pro Val Asp Ser Leu Glu Ser Thr His Gly Tyr Leu Ser Asp	340	345	350
Val Ser Ser Pro Pro Leu Met Thr Ser Pro Phe Gln Gln S r Pro Ser	355	360	365

-116-

Met Pro Leu Asn His Leu Thr Ser Met Pro Glu Ser Gln Leu Gly Met  
 370 375 380

Asn His Ile Asn Met Ala Thr Lys Gln Glu Met Ala Ala Gly Ser Asn  
 385 390 395 400

Arg Met Ala Phe Asp Ala Met Val Pro Arg Leu Thr His Leu Asn Ala  
 405 410 415

Ser Ser Pro Asn Thr Ile Met Ser Asn Gly Ser Met His Phe Thr Val  
 420 425 430

Gly Gly Ala Pro Thr Met Asn Ser Gln Cys Asp Trp Leu Ala Arg Leu  
 435 440 445

Gln Asn Gly Met Val Gln Asn Gln Tyr Asp Pro Ile Arg Asn Gly Ile  
 450 455 460

Gln Gln Gly Asn Ala Gln Gln Ala Gln Ala Leu Gln His Gly Leu Met  
 465 470 475 480

Thr Ser Leu His Asn Gly Leu Pro Ala Thr Thr Leu Ser Gln Met Met  
 485 490 495

Thr Tyr Gln Ala Met Pro Asn Thr Arg Leu Ala Asn Gln Pro His Leu  
 500 505 510

Met Gln Ala Gln Gln Met Gln Gln Gln Gln Asn Leu Gln Leu His Gln  
 515 520 525

Ser Met Gln Gln Gln His His Asn Ser Ser Thr Thr Ser Thr His Ile  
 530 535 540

Asn Ser Pro Phe Cys Ser Ser Asp Ile Ser Gln Thr Asp Leu Gln Gln  
 545 550 555 560

Met Ser Ser Asn Asn Ile His Ser Val Met Pro Gln Asp Thr Gln Ile  
 565 570 575

Phe Ala Ala Ser Leu Pro Ser Asn Leu Thr Gln Ser Met Thr Thr Ala  
 580 585 590

Gln Phe Leu Thr Pro Pro Ser Gln His Ser Tyr Ser Ser Pro Met Asp  
 595 600 605

Asn Thr Pro Ser His Gln Leu Gln Val Pro Asp His Pro Phe Leu Thr  
 610 615 620

Pro Ser Pro Glu Ser Pro Asp Gln Trp Ser Ser Ser Ser Pro His Ser  
 625 630 635 640

Asn Met Ser Asp Trp Ser Glu Gly Ile Ser Ser Pro Pro Thr  
 645 650

## (2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 666 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

-117-

Thr Pro Pro Gln Gly Glu Val Asp Ala Asp Cys Met Asp Val Asn Val  
 1 5 10 15  
 Arg Gly Pro Asp Gly Phe Thr Pro Leu Met Ile Ala Ser Cys Ser Gly  
 20 25 30  
 Gly Gly Leu Glu Thr Gly Asn Ser Glu Glu Glu Glu Asp Ala Pro Ala  
 35 40 45  
 Val Ile Ser Asp Phe Ile Tyr Gln Gly Ala Ser Leu His Asn Gln Thr  
 50 55 60  
 Asp Arg Thr Gly Glu Thr Ala Leu His Leu Ala Ala Arg Tyr Ser Arg  
 65 70 75 80  
 Ser Asp Ala Ala Lys Arg Leu Leu Glu Ala Ser Ala Asp Ala Asn Ile  
 85 90 95  
 Gln Asp Asn Met Gly Arg Thr Pro Leu His Ala Ala Val Ser Ala Asp  
 100 105 110  
 Ala Gln Gly Val Phe Gln Ile Leu Leu Arg Asn Arg Ala Thr Asp Leu  
 115 120 125  
 Asp Ala Arg Met His Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala Arg  
 130 135 140  
 Leu Ala Val Glu Gly Met Leu Glu Asp Leu Ile Asn Ser His Ala Asp  
 145 150 155 160  
 Val Asn Ala Val Asp Asp Leu Gly Lys Ser Ala Leu His Trp Ala Ala  
 165 170 175  
 Ala Val Asn Asn Val Asp Ala Ala Val Val Leu Leu Lys Asn Gly Ala  
 180 185 190  
 Asn Lys Asp Met Gln Asn Asn Lys Glu Glu Thr Pro Leu Phe Leu Ala  
 195 200 205  
 Ala Arg Glu Gly Ser Tyr Glu Thr Ala Lys Val Leu Leu Asp His Phe  
 210 215 220  
 Ala Asn Arg Asp Ile Thr Asp His Met Asp Arg Leu Pro Arg Asp Ile  
 225 230 235 240  
 Ala Gln Glu Arg Met His His Asp Ile Val Arg Leu Leu Asp Glu Tyr  
 245 250 255  
 Asn Leu Val Arg Ser Pro Gln Leu His Gly Thr Ala Leu Gly Gly Thr  
 260 265 270  
 Pro Thr Leu Ser Pro Thr Leu Cys Ser Pro Asn Gly Tyr Leu Gly Asn  
 275 280 285  
 Leu Lys Ser Ala Thr Gln Gly Lys Lys Ala Arg Lys Pro Ser Thr Lys  
 290 295 300  
 Gly Leu Ala Cys Ser Ser Lys Glu Ala Lys Asp Leu Lys Ala Arg Arg  
 305 310 315 320  
 Lys Lys Ser Gln Asp Gly Lys Gly Cys Leu Leu Asp Ser Ser Ser Met  
 325 330 335  
 Leu Ser Pro Val Asp Ser Leu Glu Ser Pro His Gly Tyr Leu Ser Asp  
 340 345 350  
 Val Ala Ser Pro Pro Leu Pro S r Pro Phe Gln Gln Ser Pro Ser Met

-118-

355	360	365
Pro Leu Ser His Leu Pro Gly Met Pro Asp Thr His Leu Gly Ile Ser		
370	375	380
His Leu Asn Val Ala Ala Lys Pro Glu Met Ala Ala Leu Ala Gly Gly		
385	390	395
Ser Arg Leu Ala Phe Glu Pro Pro Pro Pro Arg Leu Ser His Leu Pro		
	405	410
Val Ala Ser Ser Ala Ser Thr Val Leu Ser Thr Asn Gly Thr Gly Ala		
	420	425
Met Asn Phe Thr Val Gly Ala Pro Ala Ser Leu Asn Gly Gln Cys Glu		
	435	440
Trp Leu Pro Arg Leu Gln Asn Gly Met Val Pro Ser Gln Tyr Asn Pro		
	450	455
Leu Arg Pro Gly Val Thr Pro Gly Thr Leu Ser Thr Gln Ala Ala Gly		
	465	470
Leu Gln His Gly Met Met Ser Pro Ile His Ser Ser Leu Ser Thr Asn		
	485	490
Thr Leu Ser Pro Ile Ile Tyr Gln Gly Leu Pro Asn Thr Arg Leu Ala		
	500	505
Thr Gln Pro His Leu Val Gln Thr Gln Gln Val Gln Pro Gln Asn Leu		
	515	520
Gln Ile Gln Pro Gln Asn Leu Gln Pro Pro Ser Gln Pro His Leu Ser		
	530	535
Val Ser Ser Ala Ala Asn Gly His Leu Gly Arg Ser Phe Leu Ser Gly		
	545	550
Glu Pro Ser Gln Ala Asp Val Gln Pro Leu Gly Pro Ser Ser Leu Pro		
	565	570
Val His Thr Ile Leu Pro Gln Glu Ser Gln Ala Leu Pro Thr Ser Leu		
	580	585
Pro Ser Ser Met Val Pro Pro Met Thr Thr Thr Gln Phe Leu Thr Pro		
	595	600
Pro Ser Gln His Ser Tyr Ser Ser Ser Pro Val Asp Asn Thr Pro Ser		
	610	615
His Gln Leu Gln Val Pro Glu His Pro Phe Leu Thr Pro Ser Pro Glu		
	625	630
Ser Pro Asp Gln Trp Ser Ser Ser Ser Arg His Ser Asn Ile Ser Asp		
	645	650
Trp Ser Glu Gly Ile Ser Ser Pro Pro Thr		
	660	665

## (2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 681 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: unknown

-119-

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```

Thr  Pro  Pro  Gln  Gly  Glu  Val  Asp  Ala  Asp  Cys  Met  Asp  Val  Asn  Val
1      5      10     15
Arg  Gly  Pro  Asp  Gly  Phe  Thr  Pro  Leu  Met  Ile  Ala  Ser  Cys  Ser  Gly
20     25     30
Gly  Gly  Leu  Glu  Thr  Gly  Asn  Ser  Glu  Glu  Glu  Glu  Asp  Ala  Pro  Ala
35     40     45
Val  Ile  Ser  Asp  Phe  Ile  Tyr  Gln  Gly  Ala  Ser  Leu  His  Asn  Gln  Thr
50     55     60
Asp  Arg  Thr  Gly  Glu  Thr  Ala  Leu  His  Leu  Ala  Ala  Arg  Tyr  Ser  Arg
65     70     75     80
Ser  Asp  Ala  Ala  Lys  Arg  Leu  Leu  Glu  Ala  Ser  Ala  Asp  Ala  Asn  Ile
85     90     95
Gln  Asp  Asn  Met  Gly  Arg  Thr  Pro  Leu  His  Ala  Ala  Val  Ser  Ala  Asp
100    105    110
Ala  Gln  Gly  Val  Phe  Gln  Ile  Leu  Ile  Arg  Asn  Arg  Ala  Thr  Asp  Leu
115    120    125
Asp  Ala  Arg  Met  His  Asp  Gly  Thr  Thr  Pro  Leu  Ile  Leu  Ala  Ala  Arg
130    135    140
Leu  Ala  Val  Glu  Gly  Met  Leu  Glu  Asp  Leu  Ile  Asn  Ser  His  Ala  Asp
145    150    155    160
Val  Asn  Ala  Val  Asp  Asp  Leu  Gly  Lys  Ser  Ala  Leu  His  Trp  Ala  Ala
165    170    175
Ala  Val  Asn  Asn  Val  Asp  Ala  Ala  Val  Val  Leu  Leu  Lys  Asn  Gly  Ala
180    185    190
Asn  Lys  Asp  Met  Gln  Asn  Asn  Arg  Glu  Glu  Thr  Pro  Leu  Phe  Leu  Ala
195    200    205
Ala  Arg  Glu  Gly  Ser  Tyr  Glu  Thr  Ala  Lys  Val  Leu  Leu  Asp  His  Phe
210    215    220
Ala  Asn  Arg  Asp  Ile  Thr  Asp  His  Met  Asp  Arg  Leu  Pro  Arg  Asp  Ile
225    230    235    240
Ala  Gln  Glu  Arg  Met  His  His  Asp  Ile  Val  Arg  Leu  Leu  Asp  Glu  Tyr
245    250    255
Asn  Leu  Val  Arg  Ser  Pro  Gln  Leu  His  Gly  Ala  Pro  Leu  Gly  Gly  Thr
260    265    270
Pro  Thr  Leu  Ser  Pro  Pro  Leu  Cys  Ser  Pro  Asn  Gly  Tyr  Leu  Gly  Ser
275    280    285
Leu  Lys  Pro  Gly  Val  Gln  Gly  Lys  Lys  Val  Arg  Lys  Pro  Ser  Ser  Lys
290    295    300
Gly  Leu  Ala  Cys  Gly  Ser  Lys  Glu  Ala  Lys  Asp  Leu  Lys  Ala  Arg  Arg
305    310    315    320
Lys  Lys  Ser  Gln  Asp  Gly  Lys  Gly  Cys  Leu  Leu  Asp  Ser  Ser  Gly  Met

```

				325					330								335
Leu	Ser	Pro	Val	Asp	Ser	Leu	Glu	Ser	Pro	His	Gly	Tyr	Leu	Ser	Asp		
			340					345					350				
Val	Ala	Ser	Pro	Pro	Leu	Leu	Pro	Ser	Pro	Phe	Gln	Gln	Ser	Pro	Ser		
		355					360					365					
Val	Pro	Leu	Asn	His	Leu	Pro	Gly	Met	Pro	Asp	Thr	His	Leu	Gly	Ile		
	370					375					380						
Gly	His	Leu	Asn	Val	Ala	Ala	Lys	Pro	Glu	Met	Ala	Ala	Leu	Gly	Gly		
385					390					395					400		
Gly	Gly	Arg	Leu	Ala	Phe	Glu	Thr	Gly	Pro	Pro	Arg	Leu	Ser	His	Leu		
				405					410					415			
Pro	Val	Ala	Ser	Gly	Thr	Ser	Thr	Val	Leu	Gly	Ser	Ser	Ser	Gly	Gly		
			420					425						430			
Ala	Leu	Asn	Phe	Thr	Val	Gly	Gly	Ser	Thr	Ser	Leu	Asn	Gly	Gln	Cys		
	435						440					445					
Glu	Trp	Leu	Ser	Arg	Leu	Gln	Ser	Gly	Met	Val	Pro	Asn	Gln	Tyr	Asn		
450						455					460						
Pro	Leu	Arg	Gly	Ser	Val	Ala	Pro	Gly	Pro	Leu	Ser	Thr	Gln	Ala	Pro		
465					470					475					480		
Ser	Leu	Gln	His	Gly	Met	Val	Gly	Pro	Leu	His	Ser	Ser	Leu	Ala	Ala		
				485					490					495			
Ser	Ala	Leu	Ser	Gln	Met	Met	Ser	Tyr	Gln	Gly	Leu	Pro	Ser	Thr	Arg		
			500					505					510				
Leu	Ala	Thr	Gln	Pro	His	Leu	Val	Gln	Thr	Gln	Gln	Val	Gln	Pro	Gln		
		515					520					525					
Asn	Leu	Gln	Met	Gln	Gln	Gln	Asn	Leu	Gln	Pro	Ala	Asn	Ile	Gln	Gln		
530						535					540						
Gln	Gln	Ser	Leu	Gln	Pro	Pro	Pro	Pro	Pro	Pro	Gln	Pro	His	Leu	Gly		
545					550					555					560		
Val	Ser	Ser	Ala	Ala	Ser	Gly	His	Leu	Gly	Arg	Ser	Phe	Leu	Ser	Gly		
				565					570					575			
Glu	Pro	Ser	Gln	Ala	Asp	Val	Gln	Pro	Leu	Gly	Pro	Ser	Ser	Leu	Ala		
			580					585					590				
Val	His	Thr	Ile	Leu	Pro	Gln	Glu	Ser	Pro	Ala	Leu	Pro	Thr	Ser	Leu		
		595					600					605					
Pro	Ser	Ser	Leu	Val	Pro	Pro	Val	Thr	Ala	Ala	Gln	Phe	Leu	Thr	Pro		
	610					615					620						
Pro	Ser	Gln	His	Ser	Tyr	Ser	Ser	Pro	Val	Glu	Asn	Thr	Pro	Ser	His		
625					630					635					640		
Gln	Leu	Gln	Val														

-121-

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2471 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```

Met Pro Ala Leu Arg Pro Ala Leu Leu Trp Ala Leu Leu Ala Leu Trp
 1           5           10           15
Leu Cys Cys Ala Ala Pro Ala His Ala Leu Gln Cys Arg Asp Gly Tyr
          20           25           30
Glu Pro Cys Val Asn Glu Gly Met Cys Val Thr Tyr His Asn Gly Thr
          35           40           45
Gly Tyr Cys Lys Cys Pro Glu Gly Phe Leu Gly Glu Tyr Cys Gln His
          50           55           60
Arg Asp Pro Cys Glu Lys Asn Arg Cys Gln Asn Gly Gly Thr Cys Val
65           70           75           80
Ala Gln Ala Met Leu Gly Lys Ala Thr Cys Arg Cys Ala Ser Gly Phe
          85           90           95
Thr Gly Glu Asp Cys Gln Tyr Ser Thr Ser His Pro Cys Phe Val Ser
          100          105          110
Arg Pro Cys Leu Asn Gly Gly Thr Cys His Met Leu Ser Arg Asp Thr
          115          120          125
Tyr Glu Cys Thr Cys Gln Val Gly Phe Thr Gly Lys Glu Cys Gln Trp
          130          135          140
Thr Asp Ala Cys Leu Ser His Pro Cys Ala Asn Gly Ser Thr Cys Thr
          145          150          155          160
Thr Val Ala Asn Gln Phe Ser Cys Lys Cys Leu Thr Gly Phe Thr Gly
          165          170          175
Gln Lys Cys Glu Thr Asp Val Asn Glu Cys Asp Ile Pro Gly His Cys
          180          185          190
Gln His Gly Gly Thr Cys Leu Asn Leu Pro Gly Ser Tyr Gln Cys Gln
          195          200          205
Cys Pro Gln Gly Phe Thr Gly Gln Tyr Cys Asp Ser Leu Tyr Val Pro
          210          215          220
Cys Ala Pro Ser Pro Cys Val Asn Gly Gly Thr Cys Arg Gln Thr Gly
          225          230          235          240
Asp Phe Thr Phe Glu Cys Asn Cys Leu Pro Gly Phe Glu Gly Ser Thr
          245          250          255
Cys Glu Arg Asn Ile Asp Asp Cys Pro Asn His Arg Cys Gln Asn Gly
          260          265          270
Gly Val Cys Val Asp Gly Val Asn Thr Tyr Asn Cys Arg Cys Pro Pro
          275          280          285

```

-122-

Gln Trp Thr Gly Gln Phe Cys Thr Glu Asp Val Asp Glu Cys Leu Leu  
 290 295 300  
 Gln Pro Asn Ala Cys Gln Asn Gly Gly Thr Cys Ala Asn Arg Asn Gly  
 305 310 315 320  
 Gly Tyr Gly Cys Val Cys Val Asn Gly Trp Ser Gly Asp Asp Cys Ser  
 325 330 335  
 Glu Asn Ile Asp Asp Cys Ala Phe Ala Ser Cys Thr Pro Gly Ser Thr  
 340 345 350  
 Cys Ile Asp Arg Val Ala Ser Phe Ser Cys Met Cys Pro Glu Gly Lys  
 355 360 365  
 Ala Gly Leu Leu Cys His Leu Asp Asp Ala Cys Ile Ser Asn Pro Cys  
 370 375 380  
 His Lys Gly Ala Leu Cys Asp Thr Asn Pro Leu Asn Gly Gln Tyr Ile  
 385 390 395 400  
 Cys Thr Cys Pro Gln Gly Tyr Lys Gly Ala Asp Cys Thr Glu Asp Val  
 405 410 415  
 Asp Glu Cys Ala Met Ala Asn Ser Asn Pro Cys Glu His Ala Gly Lys  
 420 425 430  
 Cys Val Asn Thr Asp Gly Ala Phe His Cys Glu Cys Leu Lys Gly Tyr  
 435 440 445  
 Ala Gly Pro Arg Cys Glu Met Asp Ile Asn Glu Cys His Ser Asp Pro  
 450 455 460  
 Cys Gln Asn Asp Ala Thr Cys Leu Asp Lys Ile Gly Gly Phe Thr Cys  
 465 470 475 480  
 Leu Cys Met Pro Gly Phe Lys Gly Val His Cys Glu Leu Glu Ile Asn  
 485 490 495  
 Glu Cys Gln Ser Asn Pro Cys Val Asn Asn Gly Gln Cys Val Asp Lys  
 500 505 510  
 Val Asn Arg Phe Gln Cys Leu Cys Pro Pro Gly Phe Thr Gly Pro Val  
 515 520 525  
 Cys Gln Ile Asp Ile Asp Asp Cys Ser Ser Thr Pro Cys Leu Asn Gly  
 530 535 540  
 Ala Lys Cys Ile Asp His Pro Asn Gly Tyr Glu Cys Gln Cys Ala Thr  
 545 550 555 560  
 Gly Phe Thr Gly Val Leu Cys Glu Glu Asn Ile Asp Asn Cys Asp Pro  
 565 570 575  
 Asp Pro Cys His His Gly Gln Cys Gln Asp Gly Ile Asp Ser Tyr Thr  
 580 585 590  
 Cys Ile Cys Asn Pro Gly Tyr Met Gly Ala Ile Cys Ser Asp Gln Ile  
 595 600 605  
 Asp Glu Cys Tyr Ser Ser Pro Cys Leu Asn Asp Gly Arg Cys Il Asp  
 610 615 620  
 Leu Val Asn Gly Tyr Gln Cys Asn Cys Gln Pro Gly Thr Ser Gly Val  
 625 630 635 640  
 Asn Cys Glu Ile Asn Phe Asp Asp Cys Ala Ser Asn Pro Cys Ile His



655

ID: <WO 9407474A1 | >

-124-

Phe Ser Cys Leu Cys Pro Val Gly Phe Thr Gly Ser Phe Cys Leu His  
 1010 1015 1020  
 Glu Ile Asn Glu Cys Ser Ser His Pro Cys Leu Asn Glu Gly Thr Cys  
 1025 1030 1035 1040  
 Val Asp Gly Leu Gly Thr Tyr Arg Cys Ser Cys Pro Leu Gly Tyr Thr  
 1045 1050 1055  
 Gly Lys Asn Cys Gln Thr Leu Val Asn Leu Cys Ser Arg Ser Pro Cys  
 1060 1065 1070  
 Lys Asn Lys Gly Thr Cys Val Gln Lys Lys Ala Glu Ser Gln Cys Leu  
 1075 1080 1085  
 Cys Pro Ser Gly Trp Ala Gly Ala Tyr Cys Asp Val Pro Asn Val Ser  
 1090 1095 1100  
 Cys Asp Ile Ala Ala Ser Arg Arg Gly Val Leu Val Glu His Leu Cys  
 1105 1110 1115 1120  
 Gln His Ser Gly Val Cys Ile Asn Ala Gly Asn Thr His Tyr Cys Gln  
 1125 1130 1135  
 Cys Pro Leu Gly Tyr Thr Gly Ser Tyr Cys Glu Glu Gln Leu Asp Glu  
 1140 1145 1150  
 Cys Ala Ser Asn Pro Cys Gln His Gly Ala Thr Cys Ser Asp Phe Ile  
 1155 1160 1165  
 Gly Gly Tyr Arg Cys Glu Cys Val Pro Gly Tyr Gln Gly Val Asn Cys  
 1170 1175 1180  
 Glu Tyr Glu Val Asp Glu Cys Gln Asn Gln Pro Cys Gln Asn Gly Gly  
 1185 1190 1195 1200  
 Thr Cys Ile Asp Leu Val Asn His Phe Lys Cys Ser Cys Pro Pro Gly  
 1205 1210 1215  
 Thr Arg Gly Leu Leu Cys Glu Glu Asn Ile Asp Asp Cys Ala Arg Gly  
 1220 1225 1230  
 Pro His Cys Leu Asn Gly Gly Gln Cys Met Asp Arg Ile Gly Gly Tyr  
 1235 1240 1245  
 Ser Cys Arg Cys Leu Pro Gly Phe Ala Gly Glu Arg Cys Glu Gly Asp  
 1250 1255 1260  
 Ile Asn Glu Cys Leu Ser Asn Pro Cys Ser Ser Glu Gly Ser Leu Asp  
 1265 1270 1275 1280  
 Cys Ile Gln Leu Thr Asn Asp Tyr Leu Cys Val Cys Arg Ser Ala Phe  
 1285 1290 1295  
 Thr Gly Arg His Cys Glu Thr Phe Val Asp Val Cys Pro Gln Met Pro  
 1300 1305 1310  
 Cys Leu Asn Gly Gly Thr Cys Ala Val Ala Ser Asn Met Pro Asp Gly  
 1315 1320 1325  
 Phe Ile Cys Arg Cys Pro Pro Gly Phe Ser Gly Ala Arg Cys Gln Ser  
 1330 1335 1340  
 Ser Cys Gly Gln Val Lys Cys Arg Lys Gly Glu Gln Cys Val His Thr  
 1345 1350 1355 1360  
 Ala Ser Gly Pro Arg Cys Phe Cys Pro Ser Pro Arg Asp Cys Glu Ser

-125-

1365										1370										1375									
Gly	Cys	Ala	Ser	Ser	Pro	Cys	Gln	His	Gly	Gly	Ser	Cys	His	Pro	Gln														
			1380					1385						1390															
Arg	Gln	Pro	Pro	Tyr	Tyr	Ser	Cys	Gln	Cys	Ala	Pro	Pro	Phe	Ser	Gly														
			1395					1400						1405															
Ser	Arg	Cys	Glu	Leu	Tyr	Thr	Ala	Pro	Pro	Ser	Thr	Pro	Pro	Ala	Thr														
			1410					1415						1420															
Cys	Leu	Ser	Gln	Tyr	Cys	Ala	Asp	Lys	Ala	Arg	Asp	Gly	Val	Cys	Asp														
			1425				1430							1435															
Glu	Ala	Cys	Asn	Ser	His	Ala	Cys	Gln	Trp	Asp	Gly	Gly	Asp	Cys	Ser														
				1445					1450					1455															
Leu	Thr	Met	Glu	Asn	Pro	Trp	Ala	Asn	Cys	Ser	Ser	Pro	Leu	Pro	Cys														
			1460					1465						1470															
Trp	Asp	Tyr	Ile	Asn	Asn	Gln	Cys	Asp	Glu	Leu	Cys	Asn	Thr	Val	Glu														
			1475					1480						1485															
Cys	Leu	Phe	Asp	Asn	Phe	Glu	Cys	Gln	Gly	Asn	Ser	Lys	Thr	Cys	Lys														
			1490					1495						1500															
Tyr	Asp	Lys	Tyr	Cys	Ala	Asp	His	Phe	Lys	Asp	Asn	His	Cys	Asn	Gln														
			1505				1510							1515															
Gly	Cys	Asn	Ser	Glu	Glu	Cys	Gly	Trp	Asp	Gly	Leu	Asp	Cys	Ala	Ala														
				1525					1530					1535															
Asp	Gln	Pro	Glu	Asn	Leu	Ala	Glu	Gly	Thr	Leu	Val	Ile	Val	Val	Leu														
			1540					1545						1550															
Met	Pro	Pro	Glu	Gln	Leu	Leu	Gln	Asp	Ala	Arg	Ser	Phe	Leu	Arg	Ala														
			1555					1560						1565															
Leu	Gly	Thr	Leu	Leu	His	Thr	Asn	Leu	Arg	Ile	Lys	Arg	Asp	Ser	Gln														
			1570				1575							1580															
Gly	Glu	Leu	Met	Val	Tyr	Pro	Tyr	Tyr	Gly	Glu	Lys	Ser	Ala	Ala	Met														
			1585				1590							1595															
Lys	Lys	Gln	Arg	Met	Thr	Arg	Arg	Ser	Leu	Pro	Gly	Glu	Gln	Glu	Gln														
				1605					1610					1615															
Glu	Val	Ala	Gly	Ser	Lys	Val	Phe	Leu	Glu	Ile	Asp	Asn	Arg	Gln	Cys														
			1620					1625						1630															
Val	Gln	Asp	Ser	Asp	His	Cys	Phe	Lys	Asn	Thr	Asp	Ala	Ala	Ala	Ala														
			1635					1640						1645															
Leu	Leu	Ala	Ser	His	Ala	Ile	Gln	Gly	Thr	Leu	Ser	Tyr	Pro	Leu	Val														
			1650					1655						1660															
Ser	Val	Val	Ser	Glu	Ser	Leu	Thr	Pro	Glu	Arg	Thr	Gln	Leu	Leu	Tyr														
			1665				1670							1675															
Leu	Leu	Ala	Val	Ala	Val	Val	Ile	Ile	Leu	Phe	Ile	Ile	Leu	Leu	Gly														
				1685					1690					1695															
Val	Ile	Met	Ala	Lys	Arg	Lys	Arg	Lys	His	Gly	Ser	Leu	Trp	Leu	Pro														
			1700					1705						1710															
Glu	Gly	Phe	Thr	Leu	Arg	Arg	Asp	Ala	Ser	Asn	His	Lys	Arg	Arg	Glu														
			1715					1720						1725															

-126-

Pro Val Gly Gln Asp Ala Val Gly Leu Lys Asn Leu Ser Val Gln Val  
 1730 1735 1740  
 Ser Glu Ala Asn Leu Ile Gly Thr Gly Thr Ser Glu His Trp Val Asp  
 1745 1750 1755 1760  
 Asp Glu Gly Pro Gln Pro Lys Lys Val Lys Ala Glu Asp Glu Ala Leu  
 1765 1770 1775  
 Leu Ser Glu Glu Asp Asp Pro Ile Asp Arg Arg Pro Trp Thr Gln Gln  
 1780 1785 1790  
 His Leu Glu Ala Ala Asp Ile Arg Arg Thr Pro Ser Leu Ala Leu Thr  
 1795 1800 1805  
 Pro Pro Gln Ala Glu Gln Glu Val Asp Val Leu Asp Val Asn Val Arg  
 1810 1815 1820  
 Gly Pro Asp Gly Cys Thr Pro Leu Met Leu Ala Ser Leu Arg Gly Gly  
 1825 1830 1835 1840  
 Ser Ser Asp Leu Ser Asp Glu Asp Glu Asp Ala Glu Asp Ser Ser Ala  
 1845 1850 1855  
 Asn Ile Ile Thr Asp Leu Val Tyr Gln Gly Ala Ser Leu Gln Ala Gln  
 1860 1865 1870  
 Thr Asp Arg Thr Gly Glu Met Ala Leu His Leu Ala Ala Arg Tyr Ser  
 1875 1880 1885  
 Arg Ala Asp Ala Ala Lys Arg Leu Leu Asp Ala Gly Ala Asp Ala Asn  
 1890 1895 1900  
 Ala Gln Asp Asn Met Gly Arg Cys Pro Leu His Ala Ala Val Ala Ala  
 1905 1910 1915 1920  
 Asp Ala Gln Gly Val Phe Gln Ile Leu Ile Arg Asn Arg Val Thr Asp  
 1925 1930 1935  
 Leu Asp Ala Arg Met Asn Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala  
 1940 1945 1950  
 Arg Leu Ala Val Glu Gly Met Val Ala Glu Leu Ile Asn Cys Gln Ala  
 1955 1960 1965  
 Asp Val Asn Ala Val Asp Asp His Gly Lys Ser Ala Leu His Trp Ala  
 1970 1975 1980  
 Ala Ala Val Asn Asn Val Glu Ala Thr Leu Leu Leu Leu Lys Asn Gly  
 1985 1990 1995 2000  
 Ala Asn Arg Asp Met Gln Asp Asn Lys Glu Glu Thr Pro Leu Phe Leu  
 2005 2010 2015  
 Ala Ala Arg Glu Gly Ser Tyr Glu Ala Ala Lys Ile Leu Leu Asp His  
 2020 2025 2030  
 Phe Ala Asn Arg Asp Ile Thr Asp His Met Asp Arg Leu Pro Arg Asp  
 2035 2040 2045  
 Val Ala Arg Asp Arg Met His His Asp Ile Val Arg Leu Leu Asp Glu  
 2050 2055 2060  
 Tyr Asn Val Thr Pro Ser Pro Pro Gly Thr Val Leu Thr Ser Ala Leu  
 2065 2070 2075 2080  
 Ser Pro Val Ile Cys Gly Pro Asn Arg Ser Phe Leu Ser Leu Lys His

-127-

2085										2090										2095											
Thr	Pro	Met	Gly	Lys	Lys	Ser	Arg	Arg	Pro	Ser	Ala	Lys	Ser	Thr	Met																
			2100						2105						2110																
Pro	Thr	Ser	Leu	Pro	Asn	Leu	Ala	Lys	Glu	Ala	Lys	Asp	Ala	Lys	Gly																
			2115						2120						2125																
Ser	Arg	Arg	Lys	Lys	Ser	Leu	Ser	Glu	Lys	Val	Gln	Leu	Ser	Glu	Ser																
			2130						2135						2140																
Ser	Val	Thr	Leu	Ser	Pro	Val	Asp	Ser	Leu	Glu	Ser	Pro	His	Thr	Tyr																
			2145						2150						2155																2160
Val	Ser	Asp	Thr	Thr	Ser	Ser	Pro	Met	Ile	Thr	Ser	Pro	Gly	Ile	Leu																
									2165						2170																2175
Gln	Ala	Ser	Pro	Asn	Pro	Met	Leu	Ala	Thr	Ala	Ala	Pro	Pro	Ala	Pro																
									2180						2185																2190
Val	His	Ala	Gln	His	Ala	Leu	Ser	Phe	Ser	Asn	Leu	His	Glu	Met	Gln																
									2195						2200																2205
Pro	Leu	Ala	His	Gly	Ala	Ser	Thr	Val	Leu	Pro	Ser	Val	Ser	Gln	Leu																
									2210						2215																2220
Leu	Ser	His	His	His	Ile	Val	Ser	Pro	Gly	Ser	Gly	Ser	Ala	Gly	Ser																
									2225						2230																2235
Leu	Ser	Arg	Leu	His	Pro	Val	Pro	Val	Pro	Ala	Asp	Trp	Met	Asn	Arg																
									2245						2250																2255
Met	Glu	Val	Asn	Glu	Thr	Gln	Tyr	Asn	Glu	Met	Phe	Gly	Met	Val	Leu																
									2260						2265																2270
Ala	Pro	Ala	Glu	Gly	Thr	His	Pro	Gly	Ile	Ala	Pro	Gln	Ser	Arg	Pro																
									2275						2280																2285
Pro	Glu	Gly	Lys	His	Ile	Thr	Thr	Pro	Arg	Glu	Pro	Leu	Pro	Pro	Ile																
									2290						2295																2300
Val	Thr	Phe	Gln	Leu	Ile	Pro	Lys	Gly	Ser	Ile	Ala	Gln	Pro	Ala	Gly																
									2305						2310																2315
Ala	Pro	Gln	Pro	Gln	Ser	Thr	Cys	Pro	Pro	Ala	Val	Ala	Gly	Pro	Leu																
									2325						2330																2335
Pro	Thr	Met	Tyr	Gln	Ile	Pro	Glu	Met	Ala	Arg	Leu	Pro	Ser	Val	Ala																
									2340						2345																2350
Phe	Pro	Thr	Ala	Met	Met	Pro	Gln	Gln	Asp	Gly	Gln	Val	Ala	Gln	Thr																
									2355						2360																2365
Ile	Leu	Pro	Ala	Tyr	His	Pro	Phe	Pro	Ala	Ser	Val	Gly	Lys	Tyr	Pro																
									2370						2375																2380
Thr	Pro	Pro	Ser	Gln	His	Ser	Tyr	Ala	Ser	Ser	Asn	Ala	Ala	Glu	Arg																
									2385						2390																2395
Thr	Pro	Ser	His	Ser	Gly	His	Leu	Gln	Gly	Glu	His	Pro	Tyr	Leu	Thr																
									2405						2410																2415
Pro	Ser	Pro	Glu	Ser	Pro	Asp	Gln	Trp	Ser	Ser	Ser	Ser	Ser	Pro	His																
									2420						2425																2430
Ala	Ser	Asp	Trp	Ser	Asp	Val	Thr	Thr	Ser	Pro	Thr	Pro	Gly	Gly	Ala																
									2435						2440																2445

-128-

Gly Gly Gly Gln Arg Gly Pro Gly Thr His Met Ser Glu Pro Pro His  
 2450 2455 2460

Asn Asn Met Gln Val Tyr Ala  
 2465 2470

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2556 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met Pro Pro Leu Leu Ala Pro Leu Leu Cys Leu Ala Leu Leu Pro Ala  
 1 5 10 15

Leu Ala Ala Arg Gly Pro Arg Cys Ser Gln Pro Gly Glu Thr Cys Leu  
 20 25 30

Asn Gly Gly Lys Cys Glu Ala Ala Asn Gly Thr Glu Ala Cys Val Cys  
 35 40 45

Gly Gly Ala Phe Val Gly Pro Arg Cys Gln Asp Pro Asn Pro Cys Leu  
 50 55 60

Ser Thr Pro Cys Lys Asn Ala Gly Thr Cys His Val Val Asp Arg Arg  
 65 70 75 80

Gly Val Ala Asp Tyr Ala Cys Ser Cys Ala Leu Gly Phe Ser Gly Pro  
 85 90 95

Leu Cys Leu Thr Pro Leu Asp Asn Ala Cys Leu Thr Asn Pro Cys Arg  
 100 105 110

Asn Gly Gly Thr Cys Asp Leu Leu Thr Leu Thr Glu Tyr Lys Cys Arg  
 115 120 125

Cys Pro Pro Gly Trp Ser Gly Lys Ser Cys Gln Gln Ala Asp Pro Cys  
 130 135 140

Ala Ser Asn Pro Cys Ala Asn Gly Gly Gln Cys Leu Pro Phe Glu Ala  
 145 150 155 160

Ser Tyr Ile Cys His Cys Pro Pro Ser Phe His Gly Pro Thr Cys Arg  
 165 170 175

Gln Asp Val Asn Glu Cys Gly Gln Lys Pro Arg Leu Cys Arg His Gly  
 180 185 190

Gly Thr Cys His Asn Glu Val Gly Ser Tyr Arg Cys Val Cys Arg Ala  
 195 200 205

Thr His Thr Gly Pro Asn Cys Glu Arg Pro Tyr Val Pro Cys Ser Pro  
 210 215 220

Ser Pro Cys Gln Asn Gly Gly Thr Cys Arg Pro Thr Gly Asp Val Thr  
 225 230 235 240

His Glu Cys Ala Cys Leu Pro Gly Phe Thr Gly Gln Asn Cys Glu Glu  
 245 250 255

-129-

Asn Ile Asp Asp Cys Pro Gly Asn Asn Cys Lys Asn Gly Gly Ala Cys  
 260 265 270  
 Val Asp Gly Val Asn Thr Tyr Asn Cys Pro Cys Pro Pro Glu Trp Thr  
 275 280 285  
 Gly Gln Tyr Cys Thr Glu Asp Val Asp Glu Cys Gln Leu Met Pro Asn  
 290 295 300  
 Ala Cys Gln Asn Gly Gly Thr Cys His Asn Thr His Gly Gly Tyr Asn  
 305 310 315 320  
 Cys Val Cys Val Asn Gly Trp Thr Gly Glu Asp Cys Ser Glu Asn Ile  
 325 330 335  
 Asp Asp Cys Ala Ser Ala Ala Cys Phe His Gly Ala Thr Cys His Asp  
 340 345 350  
 Arg Val Ala Ser Phe Tyr Cys Glu Cys Pro His Gly Arg Thr Gly Leu  
 355 360 365  
 Leu Cys His Leu Asn Asp Ala Cys Ile Ser Asn Pro Cys Asn Glu Gly  
 370 375 380  
 Ser Asn Cys Asp Thr Asn Pro Val Asn Gly Lys Ala Ile Cys Thr Cys  
 385 390 395 400  
 Pro Ser Gly Tyr Thr Gly Pro Ala Cys Ser Gln Asp Val Asp Glu Cys  
 405 410 415  
 Ser Leu Gly Ala Asn Pro Cys Glu His Ala Gly Lys Cys Ile Asn Thr  
 420 425 430  
 Leu Gly Ser Phe Glu Cys Gln Cys Leu Gln Gly Tyr Thr Gly Pro Arg  
 435 440 445  
 Cys Glu Ile Asp Val Asn Glu Cys Val Ser Asn Pro Cys Gln Asn Asp  
 450 455 460  
 Ala Thr Cys Leu Asp Gln Ile Gly Glu Phe Gln Cys Met Cys Met Pro  
 465 470 475 480  
 Gly Tyr Glu Gly Val His Cys Glu Val Asn Thr Asp Glu Cys Ala Ser  
 485 490 495  
 Ser Pro Cys Leu His Asn Gly Arg Cys Leu Asp Lys Ile Asn Glu Phe  
 500 505 510  
 Gln Cys Glu Cys Pro Thr Gly Phe Thr Gly His Leu Cys Gln Tyr Asp  
 515 520 525  
 Val Asp Glu Cys Ala Ser Thr Pro Cys Lys Asn Gly Ala Lys Cys Leu  
 530 535 540  
 Asp Gly Pro Asn Thr Tyr Thr Cys Val Cys Thr Glu Gly Tyr Thr Gly  
 545 550 555 560  
 Thr His Cys Glu Val Asp Ile Asp Glu Cys Asp Pro Asp Pro Cys His  
 565 570 575  
 Tyr Gly Ser Cys Lys Asp Gly Val Ala Thr Phe Thr Cys Leu Cys Arg  
 580 585 590  
 Pro Gly Tyr Thr Gly His His Cys Glu Thr Asn Ile Asn Glu Cys Ser  
 595 600 605  
 Ser Gln Pro Cys Arg Leu Arg Gly Thr Cys Gln Asp Pro Asp Asn Ala

-130-

610					615					620					
Tyr 625	Leu	Cys	Phe	Cys	Leu 630	Lys	Gly	Thr	Thr	Gly 635	Pro	Asn	Cys	Glu	Ile 640
Asn	Leu	Asp	Asp	Cys 645	Ala	Ser	Ser	Pro	Cys 650	Asp	Ser	Gly	Thr	Cys 655	Leu
Asp	Lys	Ile	Asp 660	Gly	Tyr	Glu	Cys	Ala 665	Cys	Glu	Pro	Gly	Tyr 670	Thr	Gly
Ser	Met	Cys 675	Asn	Ser	Asn	Ile	Asp 680	Glu	Cys	Ala	Gly	Asn 685	Pro	Cys	His
Asn	Gly 690	Gly	Thr	Cys	Glu	Asp 695	Gly	Ile	Asn	Gly	Phe 700	Thr	Cys	Arg	Cys
Pro 705	Glu	Gly	Tyr	His	Asp 710	Pro	Thr	Cys	Leu	Ser 715	Glu	Val	Asn	Glu	Cys 720
Asn	Ser	Asn	Pro	Cys 725	Val	His	Gly	Ala	Cys 730	Arg	Asp	Ser	Leu	Asn 735	Gly
Tyr	Lys	Cys 740	Asp	Cys	Asp	Pro	Gly	Trp 745	Ser	Gly	Thr	Asn	Cys 750	Asp	Ile
Asn	Asn	Asn 755	Glu	Cys	Glu	Ser	Asn 760	Pro	Cys	Val	Asn	Gly 765	Gly	Thr	Cys
Lys	Asp 770	Met	Thr	Ser	Gly	Ile 775	Val	Cys	Thr	Cys	Arg 780	Glu	Gly	Phe	Ser
Gly 785	Pro	Asn	Cys	Gln	Thr 790	Asn	Ile	Asn	Glu	Cys 795	Ala	Ser	Asn	Pro	Cys 800
Leu	Asn	Lys	Gly	Thr 805	Cys	Ile	Asp	Asp	Val 810	Ala	Gly	Tyr	Lys	Cys 815	Asn
Cys	Leu	Leu	Pro 820	Tyr	Thr	Gly	Ala	Thr 825	Cys	Glu	Val	Val	Leu 830	Ala	Pro
Cys	Ala	Pro 835	Ser	Pro	Cys	Arg	Asn 840	Gly	Gly	Glu	Cys	Arg 845	Gln	Ser	Glu
Asp 850	Tyr	Glu	Ser	Phe	Ser	Cys 855	Val	Cys	Pro	Thr	Ala 860	Gly	Ala	Lys	Gly
Gln 865	Thr	Cys	Glu	Val	Asp 870	Ile	Asn	Glu	Cys	Val 875	Leu	Ser	Pro	Cys	Arg 880
His	Gly	Ala	Ser	Cys 885	Gln	Asn	Thr	His	Gly 890	Gly	Tyr	Arg	Cys	His 895	Cys
Gln	Ala	Gly	Tyr 900	Ser	Gly	Arg	Asn	Cys 905	Glu	Thr	Asp	Ile	Asp 910	Asp	Cys
Arg	Pro	Asn 915	Pro	Cys	His	Asn	Gly 920	Gly	Ser	Cys	Thr	Asp 925	Gly	Ile	Asn
Thr 930	Ala	Phe	Cys	Asp	Cys	Leu 935	Pro	Gly	Phe	Arg	Gly 940	Thr	Phe	Cys	Glu
Glu 945	Asp	Ile	Asn	Glu	Cys 950	Ala	Ser	Asp	Pro	Cys 955	Arg	Asn	Gly	Ala	Asn 960
Cys	Thr	Asp	Cys	Val 965	Asp	Ser	Tyr	Thr	Cys 970	Thr	Cys	Pro	Ala	Gly 975	Phe



-131-

Ser Gly Ile His Cys Glu Asn Asn Thr Pro Asp Cys Thr Glu Ser Ser  
 980 985 990  
 Cys Phe Asn Gly Gly Thr Cys Val Asp Gly Ile Asn Ser Phe Thr Cys  
 995 1000 1005  
 Leu Cys Pro Pro Gly Phe Thr Gly Ser Tyr Cys Gln His Val Val Asn  
 1010 1015 1020  
 Glu Cys Asp Ser Arg Pro Cys Leu Leu Gly Gly Thr Cys Gln Asp Gly  
 1025 1030 1035 1040  
 Arg Gly Leu His Arg Cys Thr Cys Pro Gln Gly Tyr Thr Gly Pro Asn  
 1045 1050 1055  
 Cys Gln Asn Leu Val His Trp Cys Asp Ser Ser Pro Cys Lys Asn Gly  
 1060 1065 1070  
 Gly Lys Cys Trp Gln Thr His Thr Gln Tyr Arg Cys Glu Cys Pro Ser  
 1075 1080 1085  
 Gly Trp Thr Gly Leu Tyr Cys Asp Val Pro Ser Val Ser Cys Glu Val  
 1090 1095 1100  
 Ala Ala Gln Arg Gln Gly Val Asp Val Ala Arg Leu Cys Gln His Gly  
 1105 1110 1115 1120  
 Gly Leu Cys Val Asp Ala Gly Asn Thr His His Cys Arg Cys Gln Ala  
 1125 1130 1135  
 Gly Tyr Thr Gly Ser Tyr Cys Glu Asp Leu Val Asp Glu Cys Ser Pro  
 1140 1145 1150  
 Ser Pro Cys Gln Asn Gly Ala Thr Cys Thr Asp Tyr Leu Gly Gly Tyr  
 1155 1160 1165  
 Ser Cys Lys Cys Val Ala Gly Tyr His Gly Val Asn Cys Ser Glu Glu  
 1170 1175 1180  
 Ile Asp Glu Cys Leu Ser His Pro Cys Gln Asn Gly Gly Thr Cys Leu  
 1185 1190 1195 1200  
 Asp Leu Pro Asn Thr Tyr Lys Cys Ser Cys Pro Arg Gly Thr Gln Gly  
 1205 1210 1215  
 Val His Cys Glu Ile Asn Val Asp Asp Cys Asn Pro Pro Val Asp Pro  
 1220 1225 1230  
 Val Ser Arg Ser Pro Lys Cys Phe Asn Asn Gly Thr Cys Val Asp Gln  
 1235 1240 1245  
 Val Gly Gly Tyr Ser Cys Thr Cys Pro Pro Gly Phe Val Gly Glu Arg  
 1250 1255 1260  
 Cys Glu Gly Asp Val Asn Glu Cys Leu Ser Asn Pro Cys Asp Ala Arg  
 1265 1270 1275 1280  
 Gly Thr Gln Asn Cys Val Gln Arg Val Asn Asp Phe His Cys Glu Cys  
 1285 1290 1295  
 Arg Ala Gly His Thr Gly Arg Arg Cys Glu Ser Val Ile Asn Gly Cys  
 1300 1305 1310  
 Lys Gly Lys Pro Cys Lys Asn Gly Gly Thr Cys Ala Val Ala Ser Asn  
 1315 1320 1325  
 Thr Ala Arg Gly Phe Ile Cys Lys Cys Pro Ala Gly Phe Glu Gly Ala

-132-

1330	1335	1340
Thr Cys Glu Asn Asp 1345	Ala Arg Thr Cys Gly 1350	Ser Leu Arg Cys Leu Asn 1355 1360
Gly Gly Thr Cys 1365	Ile Ser Gly Pro Arg 1370	Ser Pro Thr Cys Leu Cys Leu 1375
Gly Pro Phe Thr Gly 1380	Pro Glu Cys Gln Phe 1385	Pro Ala Ser Ser Pro Cys 1390
Leu Gly Gly Asn Pro Cys Tyr 1395	Asn Gln Gly Thr Cys 1400	Glu Pro Thr Ser 1405
Glu Ser Pro Phe Tyr Arg 1410	Cys Leu Cys Pro Ala 1415	Lys Phe Asn Gly Leu 1420
Leu Cys His Ile Leu Asp Tyr Ser 1425	Phe Gly Gly Ala Gly Arg 1430 1435	Asp 1440
Ile Pro Pro Pro Leu Ile Glu Glu Ala 1445	Cys Glu Leu Pro Glu Cys Gln 1450 1455	
Glu Asp Ala Gly Asn Lys Val Cys 1460	Ser Leu Gln Cys Asn Asn His Ala 1465 1470	
Cys Gly Trp Asp Gly Gly Asp 1475	Cys Ser Leu Asn Phe Asn Asp Pro Trp 1480 1485	
Lys Asn Cys Thr Gln Ser Leu Gln Cys Trp 1490 1495	Lys Tyr Phe Ser Asp Gly 1500	
His Cys Asp Ser Gln Cys Asn Ser Ala Gly 1505 1510	Cys Leu Phe Asp Gly Phe 1515 1520	
Asp Cys Gln Arg Ala Glu Gly Gln Cys 1525 1530	Asn Pro Leu Tyr Asp Gln Tyr 1535	
Cys Lys Asp His Phe Ser Asp Gly 1540 1545	His Cys Asp Gln Gly Cys Asn Ser 1550	
Ala Glu Cys Glu Trp Asp Gly Leu Asp Cys Ala 1555 1560	Glu His Val Pro Glu 1565	
Arg Leu Ala Ala Gly Thr Leu Val Val Val 1570 1575	Val Leu Met Pro Pro Glu 1580	
Gln Leu Arg Asn Ser Ser Phe His Phe Leu 1585 1590	Arg Glu Leu Ser Arg Val 1595 1600	
Leu His Thr Asn Val Val Phe Lys Arg 1605 1610	Asp Ala His Gly Gln Gln Met 1615	
Ile Phe Pro Tyr Tyr Gly Arg Glu Glu Glu 1620 1625	Leu Arg Lys His Pro Ile 1630	
Lys Arg Ala Ala Glu Gly Trp Ala Ala Pro 1635 1640	Asp Ala Leu Leu Gly Gln 1645	
Val Lys Ala Ser Leu Leu Pro Gly Gly Ser 1650 1655	Glu Gly Gly Arg Arg Arg 1660	
Arg Glu Leu Asp Pro Met Asp Val Arg 1665 1670	Gly Ser Ile Val Tyr Leu Glu 1675 1680	
Ile Asp Asn Arg Gln Cys Val Gln Ala 1685	Ser Ser Gln Cys Phe Gln Ser 1690 1695	

-133-

Ala Thr Asp Val Ala Ala Phe Leu Gly Ala Leu Ala Ser Leu Gly Ser  
 1700 1705 1710  
 Leu Asn Ile Pro Tyr Lys Ile Glu Ala Val Gln Ser Glu Thr Val Glu  
 1715 1720 1725  
 Pro Pro Pro Pro Ala Gln Leu His Phe Met Tyr Val Ala Ala Ala Ala  
 1730 1735 1740  
 Phe Val Leu Leu Phe Phe Val Gly Cys Gly Val Leu Leu Ser Arg Lys  
 1745 1750 1755 1760  
 Arg Arg Arg Gln His Gly Gln Leu Trp Phe Pro Glu Gly Phe Lys Val  
 1765 1770 1775  
 Ser Glu Ala Ser Lys Lys Lys Arg Arg Glu Glu Leu Gly Glu Asp Ser  
 1780 1785 1790  
 Val Gly Leu Lys Pro Leu Lys Asn Ala Ser Asp Gly Ala Leu Met Asp  
 1795 1800 1805  
 Asp Asn Gln Asn Glu Trp Gly Asp Glu Asp Leu Glu Thr Lys Lys Phe  
 1810 1815 1820  
 Arg Phe Glu Glu Pro Val Val Leu Pro Asp Leu Asp Asp Gln Thr Asp  
 1825 1830 1835 1840  
 His Arg Gln Trp Thr Gln Gln His Leu Asp Ala Ala Asp Leu Arg Met  
 1845 1850 1855  
 Ser Ala Met Ala Pro Thr Pro Pro Gln Gly Glu Val Asp Ala Asp Cys  
 1860 1865 1870  
 Met Asp Val Asn Val Arg Gly Pro Asp Gly Phe Thr Pro Leu Met Ile  
 1875 1880 1885  
 Ala Ser Cys Ser Gly Gly Gly Leu Glu Thr Gly Asn Ser Glu Glu Glu  
 1890 1895 1900  
 Glu Asp Ala Pro Ala Val Ile Ser Asp Phe Ile Tyr Gln Gly Ala Ser  
 1905 1910 1915 1920  
 Leu His Asn Gln Thr Asp Arg Thr Gly Glu Thr Ala Leu His Leu Ala  
 1925 1930 1935  
 Ala Arg Tyr Ser Arg Ser Asp Ala Ala Lys Arg Leu Leu Glu Ala Ser  
 1940 1945 1950  
 Ala Asp Ala Asn Ile Gln Asp Asn Met Gly Arg Thr Pro Leu His Ala  
 1955 1960 1965  
 Ala Val Ser Ala Asp Ala Gln Gly Val Phe Gln Ile Leu Ile Arg Asn  
 1970 1975 1980  
 Arg Ala Thr Asp Leu Asp Ala Arg Met His Asp Gly Thr Thr Pro Leu  
 1985 1990 1995 2000  
 Ile Leu Ala Ala Arg Leu Ala Val Glu Gly Met Leu Glu Asp Leu Ile  
 2005 2010 2015  
 Asn Ser His Ala Asp Val Asn Ala Val Asp Asp Leu Gly Lys Ser Ala  
 2020 2025 2030  
 Leu His Trp Ala Ala Ala Val Asn Asn Val Asp Ala Ala Val Val Leu  
 2035 2040 2045  
 Leu Lys Asn Gly Ala Asn Lys Asp Met Gln Asn Asn Arg Glu Glu Thr

-134-

2050	2055	2060
Pro Leu Phe Leu Ala	Ala Arg Glu Gly Ser Tyr Glu Thr Ala Lys Val	
2065	2070	2075 2080
Leu Leu Asp His Phe	Ala Asn Arg Asp Ile Thr Asp His Met Asp Arg	
	2085	2090 2095
Leu Pro Arg Asp Ile	Ala Gln Glu Arg Met His His Asp Ile Val Arg	
	2100	2105 2110
Leu Leu Asp Glu Tyr	Asn Leu Val Arg Ser Pro Gln Leu His Gly Ala	
	2115	2120 2125
Pro Leu Gly Gly Thr	Pro Thr Leu Ser Pro Pro Leu Cys Ser Pro Asn	
	2130	2135 2140
Gly Tyr Leu Gly Ser	Leu Lys Pro Gly Val Gln Gly Lys Lys Val Arg	
	2145	2150 2155 2160
Lys Pro Ser Ser Lys	Gly Leu Ala Cys Gly Ser Lys Glu Ala Lys Asp	
	2165	2170 2175
Leu Lys Ala Arg Arg	Lys Lys Ser Gln Asp Gly Lys Gly Cys Leu Leu	
	2180	2185 2190
Asp Ser Ser Gly Met	Leu Ser Pro Val Asp Ser Leu Glu Ser Pro His	
	2195	2200 2205
Gly Tyr Leu Ser Asp	Val Ala Ser Pro Pro Leu Leu Pro Ser Pro Phe	
	2210	2215 2220
Gln Gln Ser Pro Ser	Val Pro Leu Asn His Leu Pro Gly Met Pro Asp	
	2225	2230 2235 2240
Thr His Leu Gly Ile	Gly His Leu Asn Val Ala Ala Lys Pro Glu Met	
	2245	2250 2255
Ala Ala Leu Gly Gly	Gly Gly Arg Leu Ala Phe Glu Thr Gly Pro Pro	
	2260	2265 2270
Arg Leu Ser His Leu	Pro Val Ala Ser Gly Thr Ser Thr Val Leu Gly	
	2275	2280 2285
Ser Ser Ser Gly Gly	Ala Leu Asn Phe Thr Val Gly Gly Ser Thr Ser	
	2290	2295 2300
Leu Asn Gly Gln Cys	Glu Trp Leu Ser Arg Leu Gln Ser Gly Met Val	
	2305	2310 2315 2320
Pro Asn Gln Tyr Asn	Pro Leu Arg Gly Ser Val Ala Pro Gly Pro Leu	
	2325	2330 2335
Ser Thr Gln Ala Pro	Ser Leu Gln His Gly Met Val Gly Pro Leu His	
	2340	2345 2350
Ser Ser Leu Ala Ala	Ser Ala Leu Ser Gln Met Met Ser Tyr Gln Gly	
	2355	2360 2365
Leu Pro Ser Thr Arg	Leu Ala Thr Gln Pro His Leu Val Gln Thr Gln	
	2370	2375 2380
Gln Val Gln Pro Gln	Asn Leu Gln Met Gln Gln Gln Asn Leu Gln Pro	
	2385	2390 2395 2400
Ala Asn Ile Gln Gln	Gln Gln Ser Leu Gln Pro Pro Pro Pro Pro Pro	
	2405	2410 2415

-135-

Gln Pro His Leu Gly Val Ser Ser Ala Ala Ser Gly His Leu Gly Arg  
2420 2425 2430

Ser Phe Leu Ser Gly Glu Pro Ser Gln Ala Asp Val Gln Pro Leu Gly  
2435 2440 2445

Pro Ser Ser Leu Ala Val His Thr Ile Leu Pro Gln Glu Ser Pro Ala  
2450 2455 2460

Leu Pro Thr Ser Leu Pro Ser Ser Leu Val Pro Pro Val Thr Ala Ala  
2465 2470 2475 2480

Gln Phe Leu Thr Pro Pro Ser Gln His Ser Tyr Ser Ser Pro Val Glu  
2485 2490 2495

Asn Thr Pro Ser His Gln Leu Gln Val Pro Glu His Pro Phe Leu Thr  
2500 2505 2510

Pro Ser Pro Glu Ser Pro Asp Gln Trp Ser Ser Ser Ser Pro His Ser  
2515 2520 2525

Asn Val Ser Asp Trp Ser Glu Gly Val Ser Ser Pro Pro Thr Ser Met  
2530 2535 2540

Gln Ser Gln Ile Ala Arg Ile Pro Glu Ala Phe Lys  
2545 2550 2555

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9723 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 10..7419

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GGAATTCCG CCC GCC CTG CGC CCC GCT CTG CTG TGG GCG CTG CTG GCG	48
Pro Ala Leu Arg Pro Ala Leu Leu Trp Ala Leu Leu Ala	
1 5 10	
CTC TGG CTG TGC TGC GCG GCC CCC GCG CAT GCA TTG CAG TGT CGA GAT	96
Leu Trp Leu Cys Cys Ala Ala Pro Ala His Ala Leu Gln Cys Arg Asp	
15 20 25	
GGC TAT GAA CCC TGT GTA AAT GAA GGA ATG TGT GTT ACC TAC CAC AAT	144
Gly Tyr Glu Pro Cys Val Asn Glu Gly Met Cys Val Thr Tyr His Asn	
30 35 40 45	
GGC ACA GGA TAC TGC AAA TGT CCA GAA GGC TTC TTG GGG GAA TAT TGT	192
Gly Thr Gly Tyr Cys Lys Cys Pro Glu Gly Phe Leu Gly Glu Tyr Cys	
50 55 60	
CAA CAT CGA GAC CCC TGT GAG AAG AAC CGC TGC CAG AAT GGT GGG ACT	240
Gln His Arg Asp Pro Cys Glu Lys Asn Arg Cys Gln Asn Gly Gly Thr	
65 70 75	
TGT GTG GCC CAG GCC ATG CTG GGG AAA GCC ACG TGC CGA TGT GCC TCA	288
Cys Val Ala Gln Ala Met Leu Gly Lys Ala Thr Cys Arg Cys Ala Ser	

-136-

80				85				90								
GGG	TTT	ACA	GGA	GAG	GAC	TGC	CAG	TAC	TCA	ACA	TCT	CAT	CCA	TGC	TTT	336
Gly	Phe	Thr	Gly	Glu	Asp	Cys	Gln	Tyr	Ser	Thr	Ser	His	Pro	Cys	Phe	
	95					100					105					
GTG	TCT	CGA	CCC	TGC	CTG	AAT	GGC	GGC	ACA	TGC	CAT	ATG	CTC	AGC	CGG	384
Val	Ser	Arg	Pro	Cys	Leu	Asn	Gly	Gly	Thr	Cys	His	Met	Leu	Ser	Arg	
110					115					120					125	
GAT	ACC	TAT	GAG	TGC	ACC	TGT	CAA	GTC	GGG	TTT	ACA	GGT	AAG	GAG	TGC	432
Asp	Thr	Tyr	Glu	Cys	Thr	Cys	Gln	Val	Gly	Phe	Thr	Gly	Lys	Glu	Cys	
				130					135					140		
CAA	TGG	ACG	GAT	GCC	TGC	CTG	TCT	CAT	CCC	TGT	GCA	AAT	GGA	AGT	ACC	480
Gln	Trp	Thr	Asp	Ala	Cys	Leu	Ser	His	Pro	Cys	Ala	Asn	Gly	Ser	Thr	
			145					150					155			
TGT	ACC	ACT	GTG	GCC	AAC	CAG	TTC	TCC	TGC	AAA	TGC	CTC	ACA	GGC	TTC	528
Cys	Thr	Thr	Val	Ala	Asn	Gln	Phe	Ser	Cys	Lys	Cys	Leu	Thr	Gly	Phe	
		160					165					170				
ACA	GGG	CAG	AAA	TGT	GAG	ACT	GAT	GTC	AAT	GAG	TGT	GAC	ATT	CCA	GGA	576
Thr	Gly	Gln	Lys	Cys	Glu	Thr	Asp	Val	Asn	Glu	Cys	Asp	Ile	Pro	Gly	
	175					180				185						
CAC	TGC	CAG	CAT	GGT	GGC	ACC	TGC	CTC	AAC	CTG	CCT	GGT	TCC	TAC	CAG	624
His	Cys	Gln	His	Gly	Gly	Thr	Cys	Leu	Asn	Leu	Pro	Gly	Ser	Tyr	Gln	
190					195				200						205	
TGC	CAG	TGC	CCT	CAG	GGC	TTC	ACA	GGC	CAG	TAC	TGT	GAC	AGC	CTG	TAT	672
Cys	Gln	Cys	Pro	Gln	Gly	Phe	Thr	Gly	Gln	Tyr	Cys	Asp	Ser	Leu	Tyr	
				210				215						220		
GTG	CCC	TGT	GCA	CCC	TCA	CCT	TGT	GTC	AAT	GGA	GGC	ACC	TGT	CGG	CAG	720
Val	Pro	Cys	Ala	Pro	Ser	Pro	Cys	Val	Asn	Gly	Gly	Thr	Cys	Arg	Gln	
			225				230					235				
ACT	GGT	GAC	TTC	ACT	TTT	GAG	TGC	AAC	TGC	CTT	CCA	GGT	TTT	GAA	GGG	768
Thr	Gly	Asp	Phe	Thr	Phe	Glu	Cys	Asn	Cys	Leu	Pro	Gly	Phe	Glu	Gly	
		240				245						250				
AGC	ACC	TGT	GAG	AGG	AAT	ATT	GAT	GAC	TGC	CCT	AAC	CAC	AGG	TGT	CAG	816
Ser	Thr	Cys	Glu	Arg	Asn	Ile	Asp	Asp	Cys	Pro	Asn	His	Arg	Cys	Gln	
		255				260					265					
AAT	GGA	GGG	GTT	TGT	GTG	GAT	GGG	GTC	AAC	ACT	TAC	AAC	TGC	CGC	TGT	864
Asn	Gly	Gly	Val	Cys	Val	Asp	Gly	Val	Asn	Thr	Tyr	Asn	Cys	Arg	Cys	
270					275				280						285	
CCC	CCA	CAA	TGG	ACA	GGA	CAG	TTC	TGC	ACA	GAG	GAT	GTG	GAT	GAA	TGC	912
Pro	Pro	Gln	Trp	Thr	Gly	Gln	Phe	Cys	Thr	Glu	Asp	Val	Asp	Glu	Cys	
				290					295					300		
CTG	CTG	CAG	CCC	AAT	GCC	TGT	CAA	AAT	GGG	GGC	ACC	TGT	GCC	AAC	CGC	960
Leu	Leu	Gln	Pro	Asn	Ala	Cys	Gln	Asn	Gly	Gly	Thr	Cys	Ala	Asn	Arg	
			305					310					315			
AAT	GGA	GGC	TAT	GGC	TGT	GTA	TGT	GTC	AAC	GGC	TGG	AGT	GGA	GAT	GAC	1008
Asn	Gly	Gly	Tyr	Gly	Cys	Val	Cys	Val	Asn	Gly	Trp	Ser	Gly	Asp	Asp	
		320				325						330				
TGC	AGT	GAG	AAC	ATT	GAT	GAT	TGT	GCC	TTC	GCC	TCC	TGT	ACT	CCA	GGC	1056
Cys	Ser	Glu	Asn	Ile	Asp	Asp	Cys	Ala	Phe	Ala	Ser	Cys	Thr	Pro	Gly	
		335				340					345					
TCC	ACC	TGC	ATC	GAC	CGT	GTG	GCC	TCC	TTC	TCT	TGC	ATG	TGC	CCA	GAG	1104

-137-

Ser 350	Thr	Cys	Ile	Asp	Arg 355	Val	Ala	Ser	Phe	Ser 360	Cys	Met	Cys	Pro	Glu 365	
GGG	AAG	GCA	GGT	CTC	CTG	TGT	CAT	CTG	GAT	GAT	GCA	TGC	ATC	AGC	AAT	1152
Gly	Lys	Ala	Gly	Leu 370	Leu	Cys	His	Leu	Asp 375	Asp	Ala	Cys	Ile	Ser 380	Asn	
CCT	TGC	CAC	AAG	GGG	GCA	CTG	TGT	GAC	ACC	AAC	CCC	CTA	AAT	GGG	CAA	1200
Pro	Cys	His	Lys 385	Gly	Ala	Leu	Cys	Asp 390	Thr	Asn	Pro	Leu	Asn 395	Gly	Gln	
TAT	ATT	TGC	ACC	TGC	CCA	CAA	GGC	TAC	AAA	GGG	GCT	GAC	TGC	ACA	GAA	1248
Tyr	Ile	Cys 400	Thr	Cys	Pro	Gln	Gly 405	Tyr	Lys	Gly	Ala	Asp 410	Cys	Thr	Glu	
GAT	GTG	GAT	GAA	TGT	GCC	ATG	GCC	AAT	AGC	AAT	CCT	TGT	GAG	CAT	GCA	1296
Asp	Val 415	Asp	Glu	Cys	Ala	Met 420	Ala	Asn	Ser	Asn	Pro	Cys	Glu	His	Ala	
GGA	AAA	TGT	GTG	AAC	ACG	GAT	GGC	GCC	TTC	CAC	TGT	GAG	TGT	CTG	AAG	1344
Gly	Lys	Cys	Val	Asn	Thr 435	Asp	Gly	Ala	Phe	His 440	Cys	Glu	Cys	Leu	Lys 445	
GGT	TAT	GCA	GGA	CCT	CGT	TGT	GAG	ATG	GAC	ATC	AAT	GAG	TGC	CAT	TCA	1392
Gly	Tyr	Ala	Gly	Pro 450	Arg	Cys	Glu	Met	Asp 455	Ile	Asn	Glu	Cys	His	Ser 460	
GAC	CCC	TGC	CAG	AAT	GAT	GCT	ACC	TGT	CTG	GAT	AAG	ATT	GGA	GGC	TTC	1440
Asp	Pro	Cys	Gln 465	Asn	Asp	Ala	Thr	Cys 470	Leu	Asp	Lys	Ile	Gly 475	Gly	Phe	
ACA	TGT	CTG	TGC	ATG	CCA	GGT	TTC	AAA	GGT	GTG	CAT	TGT	GAA	TTA	GAA	1488
Thr	Cys 480	Leu	Cys	Met	Pro	Gly	Phe 485	Lys	Gly	Val	His	Cys 490	Glu	Leu	Glu	
ATA	AAT	GAA	TGT	CAG	AGC	AAC	CCT	TGT	GTG	AAC	AAT	GGG	CAG	TGT	GTG	1536
Ile	Asn 495	Glu	Cys	Gln	Ser	Asn 500	Pro	Cys	Val	Asn 505	Gly	Gln	Cys	Val		
GAT	AAA	GTC	AAT	CGT	TTC	CAG	TGC	CTG	TGT	CCT	CCT	GGT	TTC	ACT	GGG	1584
Asp	Lys	Val	Asn	Arg	Phe 515	Gln	Cys	Leu	Cys	Pro 520	Pro	Gly	Phe	Thr	Gly 525	
CCA	GTT	TGC	CAG	ATT	GAT	ATT	GAT	GAC	TGT	TCC	AGT	ACT	CCG	TGT	CTG	1632
Pro	Val	Cys	Gln	Ile 530	Asp	Ile	Asp	Asp	Cys 535	Ser	Ser	Thr	Pro	Cys 540	Leu	
AAT	GGG	GCA	AAG	TGT	ATC	GAT	CAC	CCG	AAT	GGC	TAT	GAA	TGC	CAG	TGT	1680
Asn	Gly	Ala	Lys 545	Cys	Ile	Asp	His	Pro 550	Asn	Gly	Tyr	Glu	Cys 555	Gln	Cys	
GCC	ACA	GGT	TTC	ACT	GGT	GTG	TTG	TGT	GAG	GAG	AAC	ATT	GAC	AAC	TGT	1728
Ala	Thr	Gly 560	Phe	Thr	Gly	Val	Leu 565	Cys	Glu	Glu	Asn 570	Ile	Asp	Asn	Cys	
GAC	CCC	GAT	CCT	TGC	CAC	CAT	GGT	CAG	TGT	CAG	GAT	GGT	ATT	GAT	TCC	1776
Asp	Pro	Asp	Pro	Cys	His	His 580	Gly	Gln	Cys	Gln	Asp 585	Gly	Ile	Asp	Ser	
TAC	ACC	TGC	ATC	TGC	AAT	CCC	GGG	TAC	ATG	GGC	GCC	ATC	TGC	AGT	GAC	1824
Tyr	Thr	Cys	Ile	Cys	Asn 595	Pro	Gly	Tyr	Met	Gly 600	Ala	Ile	Cys	Ser	Asp 605	
CAG	ATT	GAT	GAA	TGT	TAC	AGC	AGC	CCT	TGC	CTG	AAC	GAT	GGT	CGC	TGC	1872
Gln	Ile	Asp	Glu	Cys 610	Tyr	Ser	Ser	Pro	Cys 615	Leu	Asn	Asp	Gly	Arg 620	Cys	

-138-

ATT GAC CTG GTC AAT GGC TAC CAG TGC AAC TGC CAG CCA GGC ACG TCA Ile Asp Leu Val Asn Gly Tyr Gln Cys Asn Cys Gln Pro Gly Thr Ser 625 630 635	1920
GGG GTT AAT TGT GAA ATT AAT TTT GAT GAC TGT GCA AGT AAC CCT TGT Gly Val Asn Cys Glu Ile Asn Phe Asp Asp Cys Ala Ser Asn Pro Cys 640 645 650	1968
ATC CAT GGA ATC TGT ATG GAT GGC ATT AAT CGC TAC AGT TGT GTC TGC Ile His Gly Ile Cys Met Asp Gly Ile Asn Arg Tyr Ser Cys Val Cys 655 660 665	2016
TCA CCA GGA TTC ACA GGG CAG AGA TGT AAC ATT GAC ATT GAT GAG TGT Ser Pro Gly Phe Thr Gly Gln Arg Cys Asn Ile Asp Ile Asp Glu Cys 670 675 680 685	2064
GCC TCC AAT CCC TGT CGC AAG GGT GCA ACA TGT ATC AAC GGT GTG AAT Ala Ser Asn Pro Cys Arg Lys Gly Ala Thr Cys Ile Asn Gly Val Asn 690 695 700	2112
GGT TTC CGC TGT ATA TGC CCC GAG GGA CCC CAT CAC CCC AGC TGC TAC Gly Phe Arg Cys Ile Cys Pro Glu Gly Pro His His Pro Ser Cys Tyr 705 710 715	2160
TCA CAG GTG AAC GAA TGC CTG AGC AAT CCC TGC ATC CAT GGA AAC TGT Ser Gln Val Asn Glu Cys Leu Ser Asn Pro Cys Ile His Gly Asn Cys 720 725 730	2208
ACT GGA GGT CTC AGT GGA TAT AAG TGT CTC TGT GAT GCA GGC TGG GTT Thr Gly Gly Leu Ser Gly Tyr Lys Cys Leu Cys Asp Ala Gly Trp Val 735 740 745	2256
GGC ATC AAC TGT GAA GTG GAC AAA AAT GAA TGC CTT TCG AAT CCA TGC Gly Ile Asn Cys Glu Val Asp Lys Asn Glu Cys Leu Ser Asn Pro Cys 750 755 760 765	2304
CAG AAT GGA GGA ACT TGT GAC AAT CTG GTG AAT GGA TAC AGG TGT ACT Gln Asn Gly Gly Thr Cys Asp Asn Leu Val Asn Gly Tyr Arg Cys Thr 770 775 780	2352
TGC AAG AAG GGC TTT AAA GGC TAT AAC TGC CAG GTG AAT ATT GAT GAA Cys Lys Lys Gly Phe Lys Gly Tyr Asn Cys Gln Val Asn Ile Asp Glu 785 790 795	2400
TGT GCC TCA AAT CCA TGC CTG AAC CAA GGA ACC TGC TTT GAT GAC ATA Cys Ala Ser Asn Pro Cys Leu Asn Gln Gly Thr Cys Phe Asp Asp Ile 800 805 810	2448
AGT GGC TAC ACT TGC CAC TGT GTG CTG CCA TAC ACA GGC AAG AAT TGT Ser Gly Tyr Thr Cys His Cys Val Leu Pro Tyr Thr Gly Lys Asn Cys 815 820 825	2496
CAG ACA GTA TTG GCT CCC TGT TCC CCA AAC CCT TGT GAG AAT GCT GCT Gln Thr Val Leu Ala Pro Cys Ser Pro Asn Pro Cys Glu Asn Ala Ala 830 835 840 845	2544
GTT TGC AAA GAG TCA CCA AAT TTT GAG AGT TAT ACT TGC TTG TGT GCT Val Cys Lys Glu Ser Pro Asn Phe Glu Ser Tyr Thr Cys Leu Cys Ala 850 855 860	2592
CCT GGC TGG CAA GGT CAG CGG TGT ACC ATT GAC ATT GAC GAG TGT ATC Pro Gly Trp Gln Gly Gln Arg Cys Thr Ile Asp Ile Asp Glu Cys Ile 865 870 875	2640
TCC AAG CCC TGC ATG AAC CAT GGT CTC TGC CAT AAC ACC CAG GGC AGC Ser Lys Pro Cys Met Asn His Gly Leu Cys His Asn Thr Gln Gly Ser 880 885 890	2688



-139-

TAC ATG TGT GAA TGT CCA CCA GGC TTC AGT GGT ATG GAC TGT GAG GAG Tyr Met Cys Glu Cys Pro Pro Gly Phe Ser Gly Met Asp Cys Glu Glu 895 900 905	2736
GAC ATT GAT GAC TGC CTT GCC AAT CCT TGC CAG AAT GGA GGT TCC TGT Asp Ile Asp Asp Cys Leu Ala Asn Pro Cys Gln Asn Gly Gly Ser Cys 910 915 920 925	2784
ATG GAT GGA GTG AAT ACT TTC TCC TGC CTC TGC CTT CCG GGT TTC ACT Met Asp Gly Val Asn Thr Phe Ser Cys Leu Cys Leu Pro Gly Phe Thr 930 935 940	2832
GGG GAT AAG TGC CAG ACA GAC ATG AAT GAG TGT CTG AGT GAA CCC TGT Gly Asp Lys Cys Gln Thr Asp Met Asn Glu Cys Leu Ser Glu Pro Cys 945 950 955	2880
AAG AAT GGA GGG ACC TGC TCT GAC TAC GTC AAC AGT TAC ACT TGC AAG Lys Asn Gly Gly Thr Cys Ser Asp Tyr Val Asn Ser Tyr Thr Cys Lys 960 965 970	2928
TGC CAG GCA GGA TTT GAT GGA GTC CAT TGT GAG AAC AAC ATC AAT GAG Cys Gln Ala Gly Phe Asp Gly Val His Cys Glu Asn Asn Ile Asn Glu 975 980 985	2976
TGC ACT GAG AGC TCC TGT TTC AAT GGT GGC ACA TGT GTT GAT GGG ATT Cys Thr Glu Ser Ser Cys Phe Asn Gly Gly Thr Cys Val Asp Gly Ile 990 995 1000 1005	3024
AAC TCC TTC TCT TGC TTG TGC CCT GTG GGT TTC ACT GGA TCC TTC TGC Asn Ser Phe Ser Cys Leu Cys Pro Val Gly Phe Thr Gly Ser Phe Cys 1010 1015 1020	3072
CTC CAT GAG ATC AAT GAA TGC AGC TCT CAT CCA TGC CTG AAT GAG GGA Leu His Glu Ile Asn Glu Cys Ser Ser His Pro Cys Leu Asn Glu Gly 1025 1030 1035	3120
ACG TGT GTT GAT GGC CTG GGT ACC TAC CGC TGC AGC TGC CCC CTG GGC Thr Cys Val Asp Gly Leu Gly Thr Tyr Arg Cys Ser Cys Pro Leu Gly 1040 1045 1050	3168
TAC ACT GGG AAA AAC TGT CAG ACC CTG GTG AAT CTC TGC AGT CGG TCT Tyr Thr Gly Lys Asn Cys Gln Thr Leu Val Asn Leu Cys Ser Arg Ser 1055 1060 1065	3216
CCA TGT AAA AAC AAA GGT ACT TGT GTT CAG AAA AAA GCA GAG TCC CAG Pro Cys Lys Asn Lys Gly Thr Cys Val Gln Lys Lys Ala Glu Ser Gln 1070 1075 1080 1085	3264
TGC CTA TGT CCA TCT GGA TGG GCT GGT GCC TAT TGT GAC GTG CCC AAT Cys Leu Cys Pro Ser Gly Trp Ala Gly Ala Tyr Cys Asp Val Pro Asn 1090 1095 1100	3312
GTC TCT TGT GAC ATA GCA GCC TCC AGG AGA GGT GTG CTT GTT GAA CAC Val Ser Cys Asp Ile Ala Ala Ser Arg Arg Gly Val Leu Val Glu His 1105 1110 1115	3360
TTG TGC CAG CAC TCA GGT GTC TGC ATC AAT GCT GGC AAC ACG CAT TAC Leu Cys Gln His Ser Gly Val Cys Ile Asn Ala Gly Asn Thr His Tyr 1120 1125 1130	3408
TGT CAG TGC CCC CTG GGC TAT ACT GGG AGC TAC TGT GAG GAG CAA CTC Cys Gln Cys Pro Leu Gly Tyr Thr Gly S r Tyr Cys Glu Glu Gln Leu 1135 1140 1145	3456
GAT GAG TGT GCG TCC AAC CCC TGC CAG CAC GGG GCA ACA TGC AGT GAC Asp Glu Cys Ala Ser Asn Pro Cys Gln His Gly Ala Thr Cys Ser Asp 1150 1155 1160 1165	3504

-140-

TTC ATT GGT GGA TAC AGA TGC GAG TGT GTC CCA GGC TAT CAG GGT GTC Phe Ile Gly Gly Tyr Arg Cys Glu Cys Val Pro Gly Tyr Gln Gly Val 1170 1175 1180	3552
AAC TGT GAG TAT GAA GTG GAT GAG TGC CAG AAT CAG CCC TGC CAG AAT Asn Cys Glu Tyr Glu Val Asp Glu Cys Gln Asn Gln Pro Cys Gln Asn 1185 1190 1195	3600
GGA GGC ACC TGT ATT GAC CTT GTG AAC CAT TTC AAG TGC TCT TGC CCA Gly Gly Thr Cys Ile Asp Leu Val Asn His Phe Lys Cys Ser Cys Pro 1200 1205 1210	3648
CCA GGC ACT CGG GGC CTA CTC TGT GAA GAG AAC ATT GAT GAC TGT GCC Pro Gly Thr Arg Gly Leu Leu Cys Glu Glu Asn Ile Asp Asp Cys Ala 1215 1220 1225	3696
CGG GGT CCC CAT TGC CTT AAT GGT GGT CAG TGC ATG GAT AGG ATT GGA Arg Gly Pro His Cys Leu Asn Gly Gly Gln Cys Met Asp Arg Ile Gly 1230 1235 1240 1245	3744
GGC TAC AGT TGT CGC TGC TTG CCT GGC TTT GCT GGG GAG CGT TGT GAG Gly Tyr Ser Cys Arg Cys Leu Pro Gly Phe Ala Gly Glu Arg Cys Glu 1250 1255 1260	3792
GGA GAC ATC AAC GAG TGC CTC TCC AAC CCC TGC AGC TCT GAG GGC AGC Gly Asp Ile Asn Glu Cys Leu Ser Asn Pro Cys Ser Ser Glu Gly Ser 1265 1270 1275	3840
CTG GAC TGT ATA CAG CTC ACC AAT GAC TAC CTG TGT GTT TGC CGT AGT Leu Asp Cys Ile Gln Leu Thr Asn Asp Tyr Leu Cys Val Cys Arg Ser 1280 1285 1290	3888
GCC TTT ACT GGC CGG CAC TGT GAA ACC TTC GTC GAT GTG TGT CCC CAG Ala Phe Thr Gly Arg His Cys Glu Thr Phe Val Asp Val Cys Pro Gln 1295 1300 1305	3936
ATG CCC TGC CTG AAT GGA GGG ACT TGT GCT GTG GCC AGT AAC ATG CCT Met Pro Cys Leu Asn Gly Gly Thr Cys Ala Val Ala Ser Asn Met Pro 1310 1315 1320 1325	3984
GAT GGT TTC ATT TGC CGT TGT CCC CCG GGA TTT TCC GGG GCA AGG TGC Asp Gly Phe Ile Cys Arg Cys Pro Pro Gly Phe Ser Gly Ala Arg Cys 1330 1335 1340	4032
CAG AGC AGC TGT GGA CAA GTG AAA TGT AGG AAG GGG GAG CAG TGT GTG Gln Ser Ser Cys Gly Gln Val Lys Cys Arg Lys Gly Glu Gln Cys Val 1345 1350 1355	4080
CAC ACC GCC TCT GGA CCC CGC TGC TTC TGC CCC AGT CCC CGG GAC TGC His Thr Ala Ser Gly Pro Arg Cys Phe Cys Pro Ser Pro Arg Asp Cys 1360 1365 1370	4128
GAG TCA GGC TGT GCC AGT AGC CCC TGC CAG CAC GGG GGC AGC TGC CAC Glu Ser Gly Cys Ala Ser Ser Pro Cys Gln His Gly Gly Ser Cys His 1375 1380 1385	4176
CCT CAG CGC CAG CCT CCT TAT TAC TCC TGC CAG TGT GCC CCA CCA TTC Pro Gln Arg Gln Pro Pro Tyr Tyr Ser Cys Gln Cys Ala Pro Pro Phe 1390 1395 1400 1405	4224
TCG GGT AGC CGC TGT GAA CTC TAC ACG GCA CCC CCC AGC ACC CCT CCT Ser Gly Ser Arg Cys Glu Leu Tyr Thr Ala Pro Pro Ser Thr Pro Pro 1410 1415 1420	4272
GCC ACC TGT CTG AGC CAG TAT TGT GCC GAC AAA GCT CGG GAT GGC GTC Ala Thr Cys Leu Ser Gln Tyr Cys Ala Asp Lys Ala Arg Asp Gly Val 1425 1430 1435	4320

-141-

TGT GAT GAG GCC TGC AAC AGC CAT GCC TGC CAG TGG GAT GGG GGT GAC Cys Asp Glu Ala Cys Asn Ser His Ala Cys Gln Trp Asp Gly Gly Asp 1440 1445 1450	4368
TGT TCT CTC ACC ATG GAG AAC CCC TGG GCC AAC TGC TCC TCC CCA CTT Cys Ser Leu Thr Met Glu Asn Pro Trp Ala Asn Cys Ser Ser Pro Leu 1455 1460 1465	4416
CCC TGC TGG GAT TAT ATC AAC AAC CAG TGT GAT GAG CTG TGC AAC ACG Pro Cys Trp Asp Tyr Ile Asn Asn Gln Cys Asp Glu Leu Cys Asn Thr 1470 1475 1480 1485	4464
GTC GAG TGC CTG TTT GAC AAC TTT GAA TGC CAG GGG AAC AGC AAG ACA Val Glu Cys Leu Phe Asp Asn Phe Glu Cys Gln Gly Asn Ser Lys Thr 1490 1495 1500	4512
TGC AAG TAT GAC AAA TAC TGT GCA GAC CAC TTC AAA GAC AAC CAC TGT Cys Lys Tyr Asp Lys Tyr Cys Ala Asp His Phe Lys Asp Asn His Cys 1505 1510 1515	4560
AAC CAG GGG TGC AAC AGT GAG GAG TGT GGT TGG GAT GGG CTG GAC TGT Asn Gln Gly Cys Asn Ser Glu Glu Cys Gly Trp Asp Gly Leu Asp Cys 1520 1525 1530	4608
GCT GCT GAC CAA CCT GAG AAC CTG GCA GAA GGT ACC CTG GTT ATT GTG Ala Ala Asp Gln Pro Glu Asn Leu Ala Glu Gly Thr Leu Val Ile Val 1535 1540 1545	4656
GTA TTG ATG CCA CCT GAA CAA CTG CTC CAG GAT GCT CGC AGC TTC TTG Val Leu Met Pro Pro Glu Gln Leu Leu Gln Asp Ala Arg Ser Phe Leu 1550 1555 1560 1565	4704
CGG GCA CTG GGT ACC CTG CTC CAC ACC AAC CTG CGC ATT AAG CGG GAC Arg Ala Leu Gly Thr Leu Leu His Thr Asn Leu Arg Ile Lys Arg Asp 1570 1575 1580	4752
TCC CAG GGG GAA CTC ATG GTG TAC CCC TAT TAT GGT GAG AAG TCA GCT Ser Gln Gly Glu Leu Met Val Tyr Pro Tyr Tyr Gly Glu Lys Ser Ala 1585 1590 1595	4800
GCT ATG AAG AAA CAG AGG ATG ACA CGC AGA TCC CTT CCT GGT GAA CAA Ala Met Lys Lys Gln Arg Met Thr Arg Arg Ser Leu Pro Gly Glu Gln 1600 1605 1610	4848
GAA CAG GAG GTG GCT GGC TCT AAA GTC TTT CTG GAA ATT GAC AAC CGC Glu Gln Glu Val Ala Gly Ser Lys Val Phe Leu Glu Ile Asp Asn Arg 1615 1620 1625	4896
CAG TGT GTT CAA GAC TCA GAC CAC TGC TTC AAG AAC ACG GAT GCA GCA Gln Cys Val Gln Asp Ser Asp His Cys Phe Lys Asn Thr Asp Ala Ala 1630 1635 1640 1645	4944
GCA GCT CTC CTG GCC TCT CAC GCC ATA CAG GGG ACC CTG TCA TAC CCT Ala Ala Leu Leu Ala Ser His Ala Ile Gln Gly Thr Leu Ser Tyr Pro 1650 1655 1660	4992
CTT GTG TCT GTC GTC AGT GAA TCC CTG ACT CCA GAA CGC ACT CAG CTC Leu Val Ser Val Val Ser Glu Ser Leu Thr Pro Glu Arg Thr Gln Leu 1665 1670 1675	5040
CTC TAT CTC CTT GCT GTT GCT GTT GTC ATC ATT CTG TTT ATT ATT CTG Leu Tyr Leu Leu Ala Val Ala Val Ile Ile Leu Phe Ile Ile Leu 1680 1685 1690	5088
CTG GGG GTA ATC ATG GCA AAA CGA AAG CGT AAG CAT GGC TCT CTC TGG Leu Gly Val Ile Met Ala Lys Arg Lys Arg Lys His Gly Ser Leu Trp 1695 1700 1705	5136

-142-

CTG CCT GAA GGT TTC ACT CTT CGC CGA GAT GCA AGC AAT CAC AAG CGT Leu Pro Glu Gly Phe Thr Leu Arg Arg Asp Ala Ser Asn His Lys Arg 1710 1715 1720 1725	5184
CGT GAG CCA GTG GGA CAG GAT GCT GTG GGC CTG AAA AAT CTC TCA GTG Arg Glu Pro Val Gly Gln Asp Ala Val Gly Leu Lys Asn Leu Ser Val 1730 1735 1740	5232
CAA GTC TCA GAA GCT AAC CTA ATT GGT ACT GGA ACA AGT GAA CAC TGG Gln Val Ser Glu Ala Asn Leu Ile Gly Thr Gly Thr Ser Glu His Trp 1745 1750 1755	5280
GTC GAT GAT GAA GGG CCC CAG CCA AAG AAA GTA AAG GCT GAA GAT GAG Val Asp Asp Glu Gly Pro Gln Pro Lys Lys Val Lys Ala Glu Asp Glu 1760 1765 1770	5328
GCC TTA CTC TCA GAA GAA GAT GAC CCC ATT GAT CGA CGG CCA TGG ACA Ala Leu Leu Ser Glu Glu Asp Asp Pro Ile Asp Arg Arg Pro Trp Thr 1775 1780 1785	5376
CAG CAG CAC CTT GAA GCT GCA GAC ATC CGT AGG ACA CCA TCG CTG GCT Gln Gln His Leu Glu Ala Ala Asp Ile Arg Arg Thr Pro Ser Leu Ala 1790 1795 1800 1805	5424
CTC ACC CCT CCT CAG GCA GAG CAG GAG GTG GAT GTG TTA GAT GTG AAT Leu Thr Pro Pro Gln Ala Glu Gln Glu Val Asp Val Leu Asp Val Asn 1810 1815 1820	5472
GTC CGT GGC CCA GAT GGC TGC ACC CCA TTG ATG TTG GCT TCT CTC CGA Val Arg Gly Pro Asp Gly Cys Thr Pro Leu Met Leu Ala Ser Leu Arg 1825 1830 1835	5520
GGA GGC AGC TCA GAT TTG AGT GAT GAA GAT GAA GAT GCA GAG GAC TCT Gly Gly Ser Ser Asp Leu Ser Asp Glu Asp Glu Asp Ala Glu Asp Ser 1840 1845 1850	5568
TCT GCT AAC ATC ATC ACA GAC TTG GTC TAC CAG GGT GCC AGC CTC CAG Ser Ala Asn Ile Ile Thr Asp Leu Val Tyr Gln Gly Ala Ser Leu Gln 1855 1860 1865	5616
GCC CAG ACA GAC CGG ACT GGT GAG ATG GCC CTG CAC CTT GCA GCC CGC Ala Gln Thr Asp Arg Thr Gly Glu Met Ala Leu His Leu Ala Ala Arg 1870 1875 1880 1885	5664
TAC TCA CGG GCT GAT GCT GCC AAG CGT CTC CTG GAT GCA GGT GCA GAT Tyr Ser Arg Ala Asp Ala Ala Lys Arg Leu Leu Asp Ala Gly Ala Asp 1890 1895 1900	5712
GCC AAT GCC CAG GAC AAC ATG GGC CGC TGT CCA CTC CAT GCT GCA GTG Ala Asn Ala Gln Asp Asn Met Gly Arg Cys Pro Leu His Ala Ala Val 1905 1910 1915	5760
GCA GCT GAT GCC CAA GGT GTC TTC CAG ATT CTG ATT CGC AAC CGA GTA Ala Ala Asp Ala Gln Gly Val Phe Gln Ile Leu Ile Arg Asn Arg Val 1920 1925 1930	5808
ACT GAT CTA GAT GCC AGG ATG AAT GAT GGT ACT ACA CCC CTG ATC CTG Thr Asp Leu Asp Ala Arg Met Asn Asp Gly Thr Thr Pro Leu Ile Leu 1935 1940 1945	5856
GCT GCC CGC CTG GCT GTG GAG GGA ATG GTG GCA GAA CTG ATC AAC TGC Ala Ala Arg Leu Ala Val Glu Gly Met Val Ala Glu Leu Ile Asn Cys 1950 1955 1960 1965	5904
CAA GCG GAT GTG AAT GCA GTG GAT GAC CAT GGA AAA TCT GCT CTT CAC Gln Ala Asp Val Asn Ala Val Asp Asp His Gly Lys Ser Ala Leu His 1970 1975 1980	5952

-143-

TGG	GCA	GCT	GCT	GTC	AAT	AAT	GTG	GAG	GCA	ACT	CTT	TTG	TTG	TTG	AAA	6000
Trp	Ala	Ala	Ala	Val	Asn	Asn	Val	Glu	Ala	Thr	Leu	Leu	Leu	Leu	Lys	
			1985					1990								
AAT	GGG	GCC	AAC	CGA	GAC	ATG	CAG	GAC	AAC	AAG	GAA	GAG	ACA	CCT	CTG	6048
Asn	Gly	Ala	Asn	Arg	Asp	Met	Gln	Asp	Asn	Lys	Glu	Glu	Thr	Pro	Leu	
		2000					2005					2010				
TTT	CTT	GCT	GCC	CGG	GAG	GGG	AGC	TAT	GAA	GCA	GCC	AAG	ATC	CTG	TTA	6096
Phe	Leu	Ala	Ala	Arg	Glu	Gly	Ser	Tyr	Glu	Ala	Ala	Lys	Ile	Leu	Leu	
	2015					2020						2025				
GAC	CAT	TTT	GCC	AAT	CGA	GAC	ATC	ACA	GAC	CAT	ATG	GAT	CGT	CTT	CCC	6144
Asp	His	Phe	Ala	Asn	Arg	Asp	Ile	Thr	Asp	His	Met	Asp	Arg	Leu	Pro	
		2030			2035					2040					2045	
CGG	GAT	GTG	GCT	CGG	GAT	CGC	ATG	CAC	CAT	GAC	ATT	GTG	CGC	CTT	CTG	6192
Arg	Asp	Val	Ala	Arg	Asp	Arg	Met	His	His	Asp	Ile	Val	Arg	Leu	Leu	
				2050						2055				2060		
GAT	GAA	TAC	AAT	GTG	ACC	CCA	AGC	CCT	CCA	GGC	ACC	GTG	TTG	ACT	TCT	6240
Asp	Glu	Tyr	Asn	Val	Thr	Pro	Ser	Pro	Pro	Gly	Thr	Val	Leu	Thr	Ser	
			2065					2070					2075			
GCT	CTC	TCA	CCT	GTC	ATC	TGT	GGG	CCC	AAC	AGA	TCT	TTC	CTC	AGC	CTG	6288
Ala	Leu	Ser	Pro	Val	Ile	Cys	Gly	Pro	Asn	Arg	Ser	Phe	Leu	Ser	Leu	
		2080					2085					2090				
AAG	CAC	ACC	CCA	ATG	GGC	AAG	AAG	TCT	AGA	CGG	CCC	AGT	GCC	AAG	AGT	6336
Lys	His	Thr	Pro	Met	Gly	Lys	Lys	Ser	Arg	Arg	Pro	Ser	Ala	Lys	Ser	
		2095				2100					2105					
ACC	ATG	CCT	ACT	AGC	CTC	CCT	AAC	CTT	GCC	AAG	GAG	GCA	AAG	GAT	GCC	6384
Thr	Met	Pro	Thr	Ser	Leu	Pro	Asn	Leu	Ala	Lys	Glu	Ala	Lys	Asp	Ala	
		2110			2115					2120					2125	
AAG	GGT	AGT	AGG	AGG	AAG	AAG	TCT	CTG	AGT	GAG	AAG	GTC	CAA	CTG	TCT	6432
Lys	Gly	Ser	Arg	Arg	Lys	Lys	Ser	Leu	Ser	Glu	Lys	Val	Gln	Leu	Ser	
			2130					2135						2140		
GAG	AGT	TCA	GTA	ACT	TTA	TCC	CCT	GTT	GAT	TCC	CTA	GAA	TCT	CCT	CAC	6480
Glu	Ser	Ser	Val	Thr	Leu	Ser	Pro	Val	Asp	Ser	Leu	Glu	Ser	Pro	His	
			2145					2150					2155			
ACG	TAT	GTT	TCC	GAC	ACC	ACA	TCC	TCT	CCA	ATG	ATT	ACA	TCC	CCT	GGG	6528
Thr	Tyr	Val	Ser	Asp	Thr	Thr	Ser	Ser	Pro	Met	Ile	Thr	Ser	Pro	Gly	
		2160					2165					2170				
ATC	TTA	CAG	GCC	TCA	CCC	AAC	CCT	ATG	TTG	GCC	ACT	GCC	GCC	CCT	CCT	6576
Ile	Leu	Gln	Ala	Ser	Pro	Asn	Pro	Met	Leu	Ala	Thr	Ala	Ala	Pro	Pro	
		2175				2180					2185					
GCC	CCA	GTC	CAT	GCC	CAG	CAT	GCA	CTA	TCT	TTT	TCT	AAC	CTT	CAT	GAA	6624
Ala	Pro	Val	His	Ala	Gln	His	Ala	Leu	Ser	Phe	Ser	Asn	Leu	His	Glu	
	2190				2195					2200					2205	
ATG	CAG	CCT	TTG	GCA	CAT	GGG	GCC	AGC	ACT	GTG	CTT	CCC	TCA	GTG	AGC	6672
Met	Gln	Pro	Leu	Ala	His	Gly	Ala	Ser	Thr	Val	Leu	Pro	Ser	Val	Ser	
			2210					2215					2220			
CAG	TTG	CTA	TCC	CAC	CAC	CAC	ATT	GTG	TCT	CCA	GGC	AGT	GGC	AGT	GCT	6720
Gln	Leu	Leu	Ser	His	His	His	Ile	Val	Ser	Pro	Gly	Ser	Gly	Ser	Ala	
			2225				2230					2235				
GGA	AGC	TTG	AGT	AGG	CTC	CAT	CCA	GTC	CCA	GTC	CCA	GCA	GAT	TGG	ATG	6768
Gly	Ser	Leu	Ser	Arg	Leu	His	Pro	Val	Pro	Val	Pro	Ala	Asp	Trp	Met	
		2240					2245					2250				

-144-

AAC CGC ATG GAG GTG AAT GAG ACC CAG TAC AAT GAG ATG TTT GGT ATG Asn Arg Met Glu Val Asn Glu Thr Gln Tyr Asn Glu Met Phe Gly Met 2255 2260 2265	6816
GTC CTG GCT CCA GCT GAG GGC ACC CAT CCT GGC ATA GCT CCC CAG AGC Val Leu Ala Pro Ala Glu Gly Thr His Pro Gly Ile Ala Pro Gln Ser 2270 2275 2280 2285	6864
AGG CCA CCT GAA GGG AAG CAC ATA ACC ACC CCT CGG GAG CCC TTG CCC Arg Pro Pro Glu Gly Lys His Ile Thr Thr Pro Arg Glu Pro Leu Pro 2290 2295 2300	6912
CCC ATT GTG ACT TTC CAG CTC ATC CCT AAA GGC AGT ATT GCC CAA CCA Pro Ile Val Thr Phe Gln Leu Ile Pro Lys Gly Ser Ile Ala Gln Pro 2305 2310 2315	6960
GCG GGG GCT CCC CAG CCT CAG TCC ACC TGC CCT CCA GCT GTT GCG GGC Ala Gly Ala Pro Gln Pro Gln Ser Thr Cys Pro Pro Ala Val Ala Gly 2320 2325 2330	7008
CCC CTG CCC ACC ATG TAC CAG ATT CCA GAA ATG GCC CGT TTG CCC AGT Pro Leu Pro Thr Met Tyr Gln Ile Pro Glu Met Ala Arg Leu Pro Ser 2335 2340 2345	7056
GTG GCT TTC CCC ACT GCC ATG ATG CCC CAG CAG GAC GGG CAG GTA GCT Val Ala Phe Pro Thr Ala Met Met Pro Gln Gln Asp Gly Gln Val Ala 2350 2355 2360 2365	7104
CAG ACC ATT CTC CCA GCC TAT CAT CCT TTC CCA GCC TCT GTG GGC AAG Gln Thr Ile Leu Pro Ala Tyr His Pro Phe Pro Ala Ser Val Gly Lys 2370 2375 2380	7152
TAC CCC ACA CCC CCT TCA CAG CAC AGT TAT GCT TCC TCA AAT GCT GCT Tyr Pro Thr Pro Pro Ser Gln His Ser Tyr Ala Ser Ser Asn Ala Ala 2385 2390 2395	7200
GAG CGA ACA CCC AGT CAC AGT GGT CAC CTC CAG GGT GAG CAT CCC TAC Glu Arg Thr Pro Ser His Ser Gly His Leu Gln Gly Glu His Pro Tyr 2400 2405 2410	7248
CTG ACA CCA TCC CCA GAG TCT CCT GAC CAG TGG TCA AGT TCA TCA CCC Leu Thr Pro Ser Pro Glu Ser Pro Asp Gln Trp Ser Ser Ser Pro 2415 2420 2425	7296
CAC TCT GCT TCT GAC TGG TCA GAT GTG ACC ACC AGC CCT ACC CCT GGG His Ser Ala Ser Asp Trp Ser Asp Val Thr Ser Pro Thr Pro Gly 2430 2435 2440 2445	7344
GGT GCT GGA GGA GGT CAG CGG GGA CCT GGG ACA CAC ATG TCT GAG CCA Gly Ala Gly Gly Gly Gln Arg Gly Pro Gly Thr His Met Ser Glu Pro 2450 2455 2460	7392
CCA CAC AAC AAC ATG CAG GTT TAT GCG TGAGAGAGTC CACCTCCAGT Pro His Asn Asn Met Gln Val Tyr Ala 2465 2470	7439
GTAGAGACAT AACTGACTTT TGTAATGCT GCTGAGGAAC AAATGAAGGT CATCCGGGAG	7499
AGAAATGAAG AAATCTCTGG AGCCAGCTTC TAGAGGTAGG AAAGAGAAGA TGTTCTTATT	7559
CAGATAATGC AAGAGAAGCA ATTCGTCACT TCACTGGGT ATCTGCAAGG CTTATTGATT	7619
ATTCTAATCT AATAAGACAA GTTTGTGGAA ATGCAAGATG AATACAAGCC TTGGGTCCAT	7679
GTTTACTCTC TTCTATTGG AGAATAAGAT GGATGCTTAT TGAAGCCCAG ACATTCTTGC	7739
AGCTTGGA CTGATTTTAAG CCCTGCAGGC TTCTGCCATA TCCATGAGAA GATTCTACAC	7799

-145-

TAGCGTCCTG	TTGGGAATTA	TGCCCTGGAA	TTCTGCCTGA	ATTGACCTAC	GCATCTCCTC	7859
CTCCTTGGAC	ATTCTTTTGT	CTTCATTTGG	TGCTTTTGGT	TTTGCACCTC	TCCGTGATTG	7919
TAGCCCTACC	AGCATGTTAT	AGGGCAAGAC	CTTTGTGCTT	TTGATCATTC	TGGCCCATGA	7979
AAGCAACTTT	GGTCTCCTTT	CCCCTCCTGT	CTTCCC GGTA	TCCCTTGGAG	TCTCACAAGG	8039
TTTACTTTGG	TATGGTTCTC	AGCACAAACC	TTTCAAGTAT	GTTGTTTCTT	TGGAAAATGG	8099
ACATACTGTA	TTGTGTTCTC	CTGCATATAT	CATTCTCTGA	GAGAGAAGGG	GAGAAGAATA	8159
CTTTTCTTCA	ACAAATTTTG	GGGGCAGGAG	ATCCCTTCAA	GAGGCTGCAC	CTTAATTTTT	8219
CTTGTCTGTG	TGCAGGTCTT	CATATAAACT	TTACCAGGAA	GAAGGGTGTG	AGTTTGTGTG	8279
TTTTCTGTGT	ATGGGCCTGG	TCAGTGTAAG	GTTTTATCCT	TGATAGTCTA	GTTACTATGA	8339
CCCTCCCCAC	TTTTTTAAAA	CCAGAAAAAG	GTTTGGAATG	TTGGAATGAC	CAAGAGACAA	8399
GTTAACTCGT	GCAAGAGCCA	GTTACCCACC	CACAGGTCCC	CCTACTTCCT	GCCAAGCATT	8459
CCATTGACTG	CCTGTATGGA	ACACATTTGT	CCCAGATCTG	AGCATTCTAG	GCCTGTTTCA	8519
CTCACTCACC	CAGCATATGA	AACTAGTCTT	AACTGTTGAG	CCTTTCCTTT	CATATCCACA	8579
GAAGACACTG	TCTCAAATGT	TGTACCCTTG	CCATTTAGGA	CTGAACTTTC	CTTAGCCCAA	8639
GGGACCCAGT	GACAGTTGTC	TTCCGTTTGT	CAGATGATCA	GTCTCTACTG	ATTATCTTGC	8699
TGCTTAAAGG	CCTGCTCACC	AATCTTTCTT	TCACACCGTG	TGGTCCGTGT	TACTGGTATA	8759
CCCAGTATGT	TCTCACTGAA	GACATGGACT	TTATATGTTC	AAGTGCAGGA	ATTGGAAAGT	8819
TGGACTTGTT	TTCTATGATC	CAAACAGCC	CTATAAGAAG	GTTGGAAAAG	GAGGAACTAT	8879
ATAGCAGCCT	TTGCTATTTT	CTGCTACCAT	TTCTTTTCCT	CTGAAGCGGC	CATGACATTC	8939
CCTTTGGCAA	CTAACGTAGA	AACTCAACAG	AACATTTTCC	TTTCCTAGAG	TCACCTTTTA	8999
GATGATAATG	GACAACTATA	GACTTGCTCA	TTGTTTCAGAC	TGATTGCCCC	TCACCTGAAT	9059
CCACTCTCTG	TATTCATGCT	CTTGGCAATT	TCTTTGACTT	TCTTTTAAGG	GCAGAAGCAT	9119
TTTAGTTAAT	TGTAGATAAA	GAATAGTTTT	CTTCCTCTTC	TCCTTGGGCC	AGTTAATAAT	9179
TGGTCCATGG	CTACACTGCA	ACTTCCGTCC	AGTGCTGTGA	TGCCCATGAC	ACCTGCAAAA	9239
TAAGTTCTGC	CTGGGCATTT	TGTAGATATT	AACAGGTGAA	TTCCCGACTC	TTTTGGTTTG	9299
AATGACAGTT	CTCATTCCTT	CTATGGCTGC	AAGTATGCAT	CAGTGCTTCC	CACTTACCTG	9359
ATTTGTCTGT	CGGTGGCCCC	ATATGGAAAC	CCTGCGTGTC	TGTTGGCATA	ATAGTTTACA	9419
AATGGTTTTT	TCAGTCCTAT	CCAAATTTAT	TGAACCAACA	AAAATAATTA	CTTCTGCCCT	9479
GAGATAAGCA	GATTAAGTTT	GTTCAATTCTC	TGCTTTATTC	TCTCCATGTG	GCAACATTCT	9539
GTCAGCCTCT	TTCATAGTGT	GCAAACATTT	TATCATTCTA	AATGGTGACT	CTCTGCCCTT	9599
GGACCCATTT	ATTATTCACA	GATGGGGAGA	ACCTATCTGC	ATGGACCCTC	ACCATCCTCT	9659
GTGCAGCACA	CACAGTGCAG	GGAGCCAGTG	GCGATGGCGA	TGACTTTCTT	CCCCTGGGAA	9719
TTCC						9723

WHAT IS CLAIMED IS:

1. A pharmaceutical composition comprising a therapeutically effective amount of a Notch protein; and a pharmaceutically acceptable carrier.
2. The composition of claim 1 in which the Notch protein is a human Notch protein.
3. A pharmaceutical composition comprising a therapeutically effective amount of a protein, said protein comprising an amino acid sequence encoded by the DNA sequence depicted in Figure 8A (SEQ ID NO:5), 8B (SEQ ID NO:6), 8C (SEQ ID NO:7), 9A (SEQ ID NO:8), or 9B (SEQ ID NO:9), which is able to be bound by an antibody to a Notch protein; and a pharmaceutically acceptable carrier.
4. A pharmaceutical composition comprising a therapeutically effective amount of a protein, said protein comprising a Notch amino acid sequence depicted in Figure 8A (SEQ ID NO:5), 8B (SEQ ID NO:6), 8C (SEQ ID NO:7), 9A (SEQ ID NO:8), or 9B (SEQ ID NO:9), which displays one or more functional activities associated with a full-length Notch protein; and a pharmaceutically acceptable carrier.
5. A pharmaceutical composition comprising a therapeutically effective amount of a protein, said protein comprising a fragment of a human Notch protein consisting essentially of the extracellular domain of the protein; and a pharmaceutically acceptable carrier.
6. A pharmaceutical composition comprising a therapeutically effective amount of a protein, said protein comprising a region of a Notch protein containing the EGF homologous repeats of the protein; and a pharmaceutically acceptable carrier.



7. A pharmaceutical composition comprising a therapeutically effective amount of a fragment of a Notch protein lacking a portion of the EGF-homologous repeats of the protein, which fragment is able to be bound by an antibody to a Notch protein; and a pharmaceutically acceptable carrier.

5

8. A pharmaceutical composition comprising a therapeutically effective amount of a protein, said protein comprising a functionally active portion of a Notch protein; and a pharmaceutically acceptable carrier.

10

9. The composition of claim 8 in which the Notch protein is a human Notch protein.

10. A pharmaceutical composition comprising a therapeutically effective amount of a chimeric protein, said chimeric protein comprising a functionally active portion of a human Notch protein joined via a peptide bond to a sequence of a protein different from the Notch protein; and a pharmaceutically acceptable carrier.

15

11. The composition of claim 10 in which the functionally active portion of the Notch protein is encoded by the human cDNA sequence contained in plasmid hN3k as deposited with the ATCC and assigned accession number 68609, or encoded by the human cDNA sequence contained in plasmid hN5k as deposited with the ATCC and assigned accession number 68611.

20

25

12. A pharmaceutical composition comprising a therapeutically effective amount of a protein, said protein comprising the amino acid sequence depicted in Figure 10 (SEQ ID NO:11); and a pharmaceutically acceptable carrier.

30

13. A pharmaceutical composition comprising a therapeutically effective amount of a protein, said protein comprising the amino acid sequence

35

depicted in Figure 11 (SEQ ID NO:13); and a pharmaceutically acceptable carrier.

5 14. A pharmaceutical composition comprising a therapeutically effective amount of a protein, said protein comprising the portion of a human Notch protein with the greatest homology to the epidermal growth factor-like repeats 11 and 12 of the *Drosophila* Notch sequence as shown in Figure 4 (SEQ ID NO:14); and a pharmaceutically acceptable carrier.

10 15. A pharmaceutical composition comprising a therapeutically effective amount of a derivative or analog of a Notch protein, which derivative or analog is characterized by the ability *in vitro*, when expressed on the surface of a first cell, to bind to a Delta protein expressed on the surface of a second cell; and a pharmaceutically acceptable carrier.

15 16. A pharmaceutical composition comprising a therapeutically effective amount of a chimeric protein, said chimeric protein comprising a Notch protein joined via a peptide bond to a protein sequence of a protein different from the Notch protein; and a pharmaceutically acceptable carrier.

20 17. A pharmaceutical composition comprising a therapeutically effective amount of a fragment of a Notch protein, which fragment is characterized by the ability *in vitro*, when expressed on the surface of a first cell, to bind to a Delta protein expressed on the surface of a second cell; and a pharmaceutically acceptable carrier.

25 18. A pharmaceutical composition comprising a therapeutically effective amount of a chimeric protein, said chimeric protein comprising a fragment of a Notch protein joined via a peptide bond to a protein sequence of a protein different from the Notch protein, which fragment is characterized by the ability *in vitro*, when expressed on the surface of a first cell, to bind to a Delta

30

protein expressed on the surface of a second cell; and a pharmaceutically acceptable carrier.

5 19. A pharmaceutical composition comprising a therapeutically effective amount of a protein, said protein comprising a derivative or analog of a Delta protein, which derivative or analog is characterized by the ability *in vitro*, when expressed on the surface of a first cell, to bind to a Notch protein expressed on the surface of a second cell; and a pharmaceutically acceptable carrier.

10 20. A pharmaceutical composition comprising a therapeutically effective amount of a chimeric protein, said chimeric protein comprising a fragment of a Delta protein joined via a peptide bond to a protein sequence of a protein different from the Delta protein, which fragment is characterized by the ability *in vitro*, when expressed on the surface of a first cell, to bind to a Notch  
15 protein expressed on the surface of a second cell; and a pharmaceutically acceptable carrier.

21. A pharmaceutical composition comprising a therapeutically effective amount of a protein, said protein comprising a derivative or analog of a  
20 Serrate protein, which derivative or analog is characterized by the ability *in vitro*, when expressed on the surface of a first cell, to bind to a Notch protein expressed on the surface of a second cell; and a pharmaceutically acceptable carrier.

22. A pharmaceutical composition comprising a therapeutically  
25 effective amount of a derivative or analog of a Notch protein, which derivative or analog is characterized by the ability *in vitro*, when expressed on the surface of a first cell, to bind to a second protein expressed on the surface of a second cell, which second protein is selected from the group consisting of a Delta protein and a Serrate protein; and a pharmaceutically acceptable carrier.

30

23. A pharmaceutical composition comprising a therapeutically effective amount of a nucleic acid encoding a Notch protein; and a pharmaceutically acceptable carrier.

5                   24. A pharmaceutical composition comprising a therapeutically effective amount of a nucleic acid encoding a functionally active portion of a human Notch protein; and a pharmaceutically acceptable carrier.

10                   25. A pharmaceutical composition comprising a therapeutically effective amount of a nucleic acid encoding the amino acid sequence depicted in Figure 10 (SEQ ID NO:11); and a pharmaceutically acceptable carrier.

15                   26. A pharmaceutical composition comprising a therapeutically effective amount of a nucleic acid encoding the amino acid sequence depicted in Figure 11 (SEQ ID NO:13); and a pharmaceutically acceptable carrier.

20                   27. A pharmaceutical composition comprising a therapeutically effective amount of a nucleic acid encoding a fragment of a Notch protein, which fragment is characterized by the ability *in vitro*, when expressed on the surface of a first cell, to bind to a Delta protein expressed on the surface of a second cell; and a pharmaceutically acceptable carrier.

25                   28. A pharmaceutical composition comprising a therapeutically effective amount of a nucleic acid encoding a chimeric protein, said chimeric protein comprising a functionally active fragment of a human Notch protein joined via a peptide bond to a protein sequence of a protein different from the Notch protein; and a pharmaceutically acceptable carrier.

30                   29. A pharmaceutical composition comprising a therapeutically effective amount of a nucleic acid encoding a fragment of a Delta protein, which fragment is characterized by the ability *in vitro*, when expressed on the surface of

a first cell, to bind to a Notch protein expressed on the surface of a second cell;  
and a pharmaceutically acceptable carrier.

30. A pharmaceutical composition comprising a therapeutically  
5 effective amount of a nucleic acid encoding a fragment of a Serrate protein, which  
fragment is characterized by the ability *in vitro*, when expressed on the surface of  
a first cell, to bind to a Notch protein expressed on the surface of a second cell;  
and a pharmaceutically acceptable carrier.

10 31. The composition of claim 24 in which the nucleic acid is a  
nucleic acid vector.

32. A pharmaceutical composition comprising a therapeutically  
effective amount of an antibody which binds to a Notch protein; and a  
15 pharmaceutically acceptable carrier.

33. A pharmaceutical composition comprising a therapeutically  
effective amount of a fragment or derivative of an antibody to a Notch protein  
containing the idiotype of the antibody; and a pharmaceutically acceptable carrier.

20 34. A method of treating or preventing a disease or disorder in a  
subject comprising administering to a subject in need of such treatment or  
prevention a therapeutically effective amount of a molecule which antagonizes the  
function of a Notch protein.

25 35. The method according to claim 34 in which the disease or  
disorder is a malignancy characterized by increased Notch activity or increased  
expression of a Notch protein or of a Notch derivative capable of being bound by  
an anti-Notch antibody, relative to said Notch activity or expression in an  
30 analogous non-malignant sample.

36. The method according to claim 34 in which the disease or disorder is cervical cancer.

5 37. The method according to claim 34 in which the disease or disorder is breast cancer.

38. The method according to claim 34 in which the disease or disorder is colon cancer.

10 39. The method according to claim 35 in which the malignancy is selected from the group consisting of melanoma, seminoma, and lung cancer.

15 40. The method according to claim 35 in which the subject is a human.

41. The method according to claim 36, 37 or 38 in which the molecule is an antibody to Notch or a portion of said antibody containing the binding domain thereof.

20 42. The method according to claim 36, 37 or 38 in which the molecule is a protein consisting of at least the extracellular domain of a Notch protein or a portion thereof capable of binding to a Notch ligand.

25 43. The method according to claim 36, 37 or 38 in which the molecule is a protein consisting of at least the EGF homologous repeats of a Notch protein.

30 44. The method according to claim 36, 37 or 38 in which the molecule is a protein consisting of at least an adhesive fragment of a Notch protein.

45. The method according to claim 36, 37 or 38 in which the molecule is an oligonucleotide which (a) consists of at least six nucleotides; (b) comprises a sequence complementary to at least a portion of an RNA transcript of a Notch gene; and (c) is hybridizable to the RNA transcript.

5

46. A method of treating or preventing a disease or disorder in a subject in need of such treatment or prevention comprising administering to the subject a therapeutically effective amount of a molecule which promotes the function of a Notch protein.

10

47. A method of treating or preventing a malignancy in a subject comprising administering to a subject in need of such treatment or prevention an effective amount of a Notch protein.

15

48. A method of treating or preventing a malignancy in a subject comprising administering to a subject in need of such treatment or prevention an effective amount of a functionally active portion of a Notch protein.

20

49. The method according to claim 47 in which the Notch protein is a human Notch protein.

25

50. A method of treating or preventing a malignancy in a subject comprising administering to a subject in need of such treatment or prevention an effective amount of a chimeric protein, said protein comprising a functionally active portion of a Notch protein joined via a peptide bond to a protein sequence of a protein different from the Notch protein.

30

51. The method according to claim 49 in which the human Notch protein comprises the amino acid sequence depicted in Figure 10 (SEQ ID NO:11) or Figure 11 (SEQ ID NO:13).

35

52. A method of treating or preventing a malignancy in a subject comprising administering to a subject in need of such treatment or prevention an effective amount of a derivative or analog of a Notch protein, which derivative or analog is characterized by the ability *in vitro*, when expressed on the surface of a first cell, to bind to a second protein expressed on the surface of a second cell, which second protein is selected from the group consisting of a Delta protein and a Serrate protein.

53. A method of treating or preventing a malignancy in a subject comprising administering to a subject in need of such treatment or prevention an effective amount of a derivative or analog of a Delta protein, which derivative or analog is characterized by the ability *in vitro*, when expressed on the surface of a first cell, to bind to a Notch protein expressed on the surface of a second cell.

54. A method of treating or preventing a malignancy in a subject comprising administering to a subject in need of such treatment or prevention an effective amount of a derivative or analog of a Serrate protein, which derivative or analog is characterized by the ability *in vitro*, when expressed on the surface of a first cell, to bind to a Notch protein expressed on the surface of a second cell.

55. A method of treating or preventing a malignancy in a subject comprising administering to a subject in need of such treatment or prevention an effective amount of a nucleic acid encoding a Notch protein.

56. A method of treating or preventing a malignancy in a subject comprising administering to a subject in need of such treatment or prevention an effective amount of a nucleic acid encoding a functionally active portion of a Notch protein.

57. The method according to claim 55 in which the subject is human and the Notch protein is a human Notch protein.



58. A method of treating or preventing a malignancy in a subject comprising administering to a subject in need of such treatment or prevention an effective amount of a nucleic acid encoding a fragment of a Notch protein, which fragment is characterized by the ability *in vitro*, when expressed on the surface of  
5 a first cell, to bind to a second protein expressed on the surface of a second cell, which second protein is selected from the group consisting of a Delta protein and a Serrate protein.

59. A method of treating or preventing a malignancy in a subject  
10 comprising administering to a subject in need of such treatment or prevention an effective amount of a nucleic acid encoding a fragment of a Delta protein, which fragment is characterized by the ability *in vitro*, when expressed on the surface of a first cell, to bind to a Notch protein expressed on the surface of a second cell.

60. A method of treating or preventing a malignancy in a subject  
15 comprising administering to a subject in need of such treatment or prevention an effective amount of a nucleic acid encoding a fragment of a Serrate protein, which fragment is characterized by the ability *in vitro*, when expressed on the surface of a first cell, to bind to a Notch protein expressed on the surface of a second cell.  
20

61. A method of treating or preventing a malignancy in a subject comprising administering to a subject in need of such treatment or prevention an effective amount of antibody to a Notch protein.

62. The method according to claim 58 in which the antibody is  
25 monoclonal.

63. A method for treating a patient with a tumor, of a tumor type characterized by expression of a Notch gene, comprising administering to the  
30 patient an effective amount of an oligonucleotide, which oligonucleotide (a) consists of at least six nucleotides; (b) comprises a sequence complementary to at

least a portion of an RNA transcript of the Notch gene; and (c) is hybridizable to the RNA transcript.

5           64. The method according to claim 60 in which the patient is a human, and the Notch gene is a human gene.

          65. An isolated oligonucleotide consisting of at least six nucleotides, and comprising a sequence complementary to at least a portion of an RNA transcript of a Notch gene, which oligonucleotide is hybridizable to the  
10       RNA transcript.

          66. A pharmaceutical composition comprising the oligonucleotide of claim 65; and a pharmaceutically acceptable carrier.

15           67. A method of inhibiting the expression of a nucleic acid sequence encoding a Notch protein in a cell comprising providing the cell with an effective amount of the oligonucleotide of claim 65.

          68. A method of diagnosing a disease or disorder characterized by  
20       an aberrant level of Notch protein or activity in a patient, comprising measuring the level of Notch protein expression or activity in a sample derived from the patient, in which an increase or decrease in Notch protein or activity in the patient sample relative to the level found in such a sample from a normal individual indicates the presence of the disease or disorder in the patient.

25           69. A method of diagnosing a malignancy characterized by an increased amount of a Notch protein or of a Notch derivative capable of being bound by an anti-Notch antibody, comprising measuring the amount of a Notch protein or of a Notch derivative capable of being bound by an anti-Notch  
30       antibody, in a sample containing or suspected of containing malignant cells from a patient, in which an increase in the amount of the Notch protein or of the Notch

derivative capable of being bound by an anti-Notch antibody, in the sample, relative to said amount found in an analogous sample of non-malignant cells indicates the presence of the disease or disorder in the patient.

5                   70. The method according to claim 69 in which the malignancy is cervical cancer.

71. The method according to claim 69 in which the malignancy is breast cancer.

10

72. The method according to claim 69 in which the malignancy is colon cancer.

15                   73. The method according to claim 69 in which the malignancy is selected from the group consisting of melanoma, seminoma, and lung cancer.

74. The method according to claim 69 in which the amount of the Notch protein or derivative is measured by a method comprising contacting the sample with an anti-Notch antibody such that immunospecific binding can occur, and measuring the amount of any immunospecific binding of the antibody that occurs.

20

75. A method of treating or preventing a nervous system disorder in a subject comprising administering to a subject in need of such treatment or prevention an effective amount of a functionally active portion of a Notch protein.

25

76. A method of promoting tissue regeneration or repair in a subject comprising administering to a subject an effective amount of a functionally active portion of a Notch protein.

30

35

77. A method of treating a benign dysproliferative disorder in a subject comprising administering to a subject in need of such treatment an effective amount of a functionally active portion of a Notch protein, in which the disorder is selected from the group consisting of cirrhosis of the liver, psoriasis,  
5 keloids, and baldness.

78. A substantially purified human Notch protein comprising the amino acid sequence encoded by the hN homolog as depicted in Figure 13 from amino acid numbers 1 through 2169 (SEQ ID NO:19).  
10

79. A substantially purified human Notch protein comprising the amino acid sequence encoded by the hN homolog as depicted in Figure 13 from amino acid numbers about 26 through 2169 (as contained in SEQ ID NO:19).

80. A substantially purified protein comprising the extracellular domain of the mature human Notch protein encoded by the hN homolog, as depicted in Figure 13 from amino acid numbers about 26 through 1677 (as contained in SEQ ID NO:19).  
15

81. A substantially purified protein comprising the EGF homologous repeats of the mature human Notch protein encoded by the hN homolog, as depicted in Figure 13 from amino acid numbers 26 through 1413 (as contained in SEQ ID NO:19).  
20

82. A substantially purified protein comprising the EGF like repeats 11 and 12 of the mature human Notch protein encoded by the hN homolog, as depicted in Figure 13 (as contained in SEQ ID NO:19).  
25

83. A substantially purified protein consisting essentially of the extracellular domain of the mature human Notch protein encoded by the hN  
30

homolog, as depicted in Figure 13 from amino acid numbers about 26 through 1677 (as contained in SEQ ID NO:19).

5                   84. A substantially purified nucleic acid encoding the protein of claim 78.

                  85. A substantially purified nucleic acid encoding the protein of claim 79.

10                  86. A substantially purified nucleic acid encoding the protein of claim 80.

                  87. A substantially purified nucleic acid encoding the protein of claim 82.  
15

                  88. The nucleic acid of claim 85 which is a DNA molecule comprising the sequence depicted in Figure 17 from nucleotide numbers 82 through 7419 (as contained in SEQ ID NO:21).

20                  89. The nucleic acid of claim 80 in which the sequence encoding the extracellular domain is as presented in Figure 17 (as contained in SEQ ID NO:21).

                  90. A recombinant cell containing the nucleic acid of claim 84, 87  
25                  or 88.

                  91. The composition of claim 2 in which the Notch protein comprises the amino acid sequence encoded by the hN homolog as depicted in Figure 13 from amino acid numbers 26 through 2169 (as contained in  
30                  SEQ ID NO:19).

92. A composition comprising a therapeutically effective amount of a Notch protein or Notch derivative, said derivative being capable of being bound by an anti-Notch antibody, for use as a medicament.

5                   93. A composition comprising a therapeutically effective amount of a molecule which antagonizes the function of a Notch protein, for use as a medicament.

10                   94. Use of a composition comprising a molecule which antagonizes the function of a Notch protein, for the manufacture of a medicament for the treatment of cervical cancer, breast cancer, or colon cancer.

15

20

25

30

35

1/68

GAATTCGGAG GAATTATTCA AAACATAAAC ACAATAAACA ATTTGAGTAG TTGCCGCACA	60
CACACACACA CACAGCCCGT GGATTATTAC ACTAAAAGCG ACACTCAATC CAAAAAATCA	120
GCAACAAAAA CATCAATAAA C ATG CAT TGG ATT AAA TGT TTA TTA ACA GCA	171
Met His Trp Ile Lys Cys Leu Leu Thr Ala	
1 5 10	
TTC ATT TGC TTC ACA GTC ATC GTG CAG GTT CAC AGT TCC GGC AGC TTT	219
Phe Ile Cys Phe Thr Val Ile Val Gln Val His Ser Ser Gly Ser Phe	
15 20 25	
GAG TTG CGC CTG AAG TAC TTC AGC AAC GAT CAC GGG CGG GAC AAC GAG	267
Glu Leu Arg Leu Lys Tyr Phe Ser Asn Asp His Gly Arg Asp Asn Glu	
30 35 40	
GGT CGC TGC TGC AGC GGG GAG TCG GAC GGA GCG ACG GGC AAG TGC CTG	315
Gly Arg Cys Cys Ser Gly Glu Ser Asp Gly Ala Thr Gly Lys Cys Leu	
45 50 55	
GGC AGC TGC AAG ACG CGG TTT CGC GTC TGC CTA AAG CAC TAC CAG GCC	363
Gly Ser Cys Lys Thr Arg Phe Arg Val Cys Leu Lys His Tyr Gln Ala	
60 65 70	
ACC ATC GAC ACC ACC TCC CAG TGC ACC TAC GGG GAC GTG ATC ACG CCC	411
Thr Ile Asp Thr Thr Ser Gln Cys Thr Tyr Gly Asp Val Ile Thr Pro	
75 80 85 90	
ATT CTC GGC GAG AAC TCG GTC AAT CTG ACC GAC GCC CAG CGC TTC CAG	459
Ile Leu Gly Glu Asn Ser Val Asn Leu Thr Asp Ala Gln Arg Phe Gln	
95 100 105	
AAC AAG GGC TTC ACG AAT CCC ATC CAG TTC CCC TTC TCG TTC TCA TGG	507
Asn Lys Gly Phe Thr Asn Pro Ile Gln Phe Pro Phe Ser Phe Ser Trp	
110 115 120	

FIG.1A

2 / 68

CCG GGT ACC TTC TCG CTG ATC GTC GAG GCC TGG CAT GAT ACG AAC AAT	555
Pro Gly Thr Phe Ser Leu Ile Val Glu Ala Trp His Asp Thr Asn Asn	
125 130 135	
AGC GGC AAT GCG CGA ACC AAC AAG CTC CTC ATC CAG CGA CTC TTG GTG	603
Ser Gly Asn Ala Arg Thr Asn Lys Leu Leu Ile Gln Arg Leu Leu Val	
140 145 150	
CAG CAG GTA CTG GAG GTG TCC TCC GAA TGG AAG ACG AAC AAG TCG GAA	651
Gln Gln Val Leu Glu Val Ser Ser Glu Trp Lys Thr Asn Lys Ser Glu	
155 160 165 170	
TCG CAG TAC ACG TCG CTG GAG TAC GAT TTC CGT GTC ACC TGC GAT CTC	699
Ser Gln Tyr Thr Ser Leu Glu Tyr Asp Phe Arg Val Thr Cys Asp Leu	
175 180 185	
AAC TAC TAC GGA TCC GGC TGT GCC AAG TTC TGC CGG CCC CGC GAC GAT	747
Asn Tyr Tyr Gly Ser Gly Cys Ala Lys Phe Cys Arg Pro Arg Asp Asp	
190 195 200	
TCA TTT GGA CAC TCG ACT TGC TCG GAG ACG GGC GAA ATT ATC TGT TTG	795
Ser Phe Gly His Ser Thr Cys Ser Glu Thr Gly Glu Ile Ile Cys Leu	
205 210 215	
ACC GGA TGG CAG GGC GAT TAC TGT CAC ATA CCC AAA TGC GCC AAA GGC	843
Thr Gly Trp Gln Gly Asp Tyr Cys His Ile Pro Lys Cys Ala Lys Gly	
220 225 230	
TGT GAA CAT GGA CAT TGC GAC AAA CCC AAT CAA TGC GTT TGC CAA CTG	891
Cys Glu His Gly His Cys Asp Lys Pro Asn Gln Cys Val Cys Gln Leu	
235 240 245 250	
GGC TGG AAG GGA GCC TTG TGC AAC GAG TGC GTT CTG GAA CCG AAC TGC	939
Gly Trp Lys Gly Ala Leu Cys Asn Glu Cys Val Leu Glu Pro Asn Cys	
255 260 265	

FIG.1B



3/68

ATC CAT GGC ACC TGC AAC AAA CCC TGG ACT TGC ATC TGC AAC GAG GGT Ile His Gly Thr Cys Asn Lys Pro Trp Thr Cys Ile Cys Asn Glu Gly 270 275 280	987
TGG GGA GGC TTG TAC TGC AAC CAG GAT CTG AAC TAC TGC ACC AAC CAC Trp Gly Gly Leu Tyr Cys Asn Gln Asp Leu Asn Tyr Cys Thr Asn His 285 290 295	1035
AGA CCC TGC AAG AAT GGC GGA ACC TGC TTC AAC ACC GGC GAG GGA TTG Arg Pro Cys Lys Asn Gly Gly Thr Cys Phe Asn Thr Gly Glu Gly Leu 300 305 310	1083
TAC ACA TGC AAA TGC GCT CCA GGA TAC AGT GGT GAT GAT TGC GAA AAT Tyr Thr Cys Lys Cys Ala Pro Gly Tyr Ser Gly Asp Asp Cys Glu Asn 315 320 325 330	1131
GAG ATC TAC TCC TGC GAT GCC GAT GTC AAT CCC TGC CAG AAT GGT GGT Glu Ile Tyr Ser Cys Asp Ala Asp Val Asn Pro Cys Gln Asn Gly Gly 335 340 345	1179
ACC TGC ATC GAT GAG CCG CAC ACA AAA ACC GGC TAC AAG TGT CAT TGC Thr Cys Ile Asp Glu Pro His Thr Lys Thr Gly Tyr Lys Cys His Cys 350 355 360	1227
GCC AAC GGC TGG AGC GGA AAG ATG TGC GAG GAG AAA GTG CTC ACG TGT Ala Asn Gly Trp Ser Gly Lys Met Cys Glu Glu Lys Val Leu Thr Cys 365 370 375	1275
TCG GAC AAA CCC TGT CAT CAG GGA ATC TGC CGC AAC GTT CGT CCT GGC Ser Asp Lys Pro Cys His Gln Gly Ile Cys Arg Asn Val Arg Pro Gly 380 385 390	1323
TTG GGA AGC AAG GGT CAG GGC TAC CAG TGC GAA TGT CCC ATT GGC TAC Leu Gly Ser Lys Gly Gln Gly Tyr Gln Cys Glu Cys Pro Ile Gly Tyr 395 400 405 410	1371

FIG.1C

4/68

AGC GGA CCC AAC TGC GAT CTC CAG CTG GAC AAC TGC AGT CCG AAT CCA	1419
Ser Gly Pro Asn Cys Asp Leu Gln Leu Asp Asn Cys Ser Pro Asn Pro	
415 420 425	
TGC ATA AAC GGT GGA AGC TGT CAG CCG AGC GGA AAG TGT ATT TGC CCA	1467
Cys Ile Asn Gly Gly Ser Cys Gln Pro Ser Gly Lys Cys Ile Cys Pro	
430 435 440	
GCG GGA TTT TCG GGA ACG AGA TGC GAG ACC AAC ATT GAC GAT TGT CTT	1515
Ala Gly Phe Ser Gly Thr Arg Cys Glu Thr Asn Ile Asp Asp Cys Leu	
445 450 455	
GGC CAC CAG TGC GAG AAC GGA GGC ACC TGC ATA GAT ATG GTC AAC CAA	1563
Gly His Gln Cys Glu Asn Gly Gly Thr Cys Ile Asp Met Val Asn Gln	
460 465 470	
TAT CGC TGC CAA TGC GTT CCC GGT TTC CAT GGC ACC CAC TGT AGT AGC	1611
Tyr Arg Cys Gln Cys Val Pro Gly Phe His Gly Thr His Cys Ser Ser	
475 480 485 490	
AAA GTT GAC TTG TGC CTC ATC AGA CCG TGT GCC AAT GGA GGA ACC TGC	1659
Lys Val Asp Leu Cys Leu Ile Arg Pro Cys Ala Asn Gly Gly Thr Cys	
495 500 505	
TTG AAT CTC AAC AAC GAT TAC CAG TGC ACC TGT CGT GCG GGA TTT ACT	1707
Leu Asn Leu Asn Asn Asp Tyr Gln Cys Thr Cys Arg Ala Gly Phe Thr	
510 515 520	
GGC AAG GAT TGC TCT GTG GAC ATC GAT GAG TGC AGC AGT GGA CCC TGT	1755
Gly Lys Asp Cys Ser Val Asp Ile Asp Glu Cys Ser Ser Gly Pro Cys	
525 530 535	
CAT AAC GGC GGC ACT TGC ATG AAC CGC GTC AAT TCG TTC GAA TGC GTG	1803
His Asn Gly Gly Thr Cys Met Asn Arg Val Asn Ser Phe Glu Cys Val	
540 550	

FIG.1D

5/68

TGT GCC AAT GGT TTC AGG GGC AAG CAG TGC GAT GAG GAG TCC TAC GAT Cys Ala Asn Gly Phe Arg Gly Lys Gln Cys Asp Glu Glu Ser Tyr Asp 555 560 565 570	1851
TCG GTG ACC TTC GAT GCC CAC CAA TAT GGA GCG ACC ACA CAA GCG AGA Ser Val Thr Phe Asp Ala His Gln Tyr Gly Ala Thr Thr Gln Ala Arg 575 580 585	1899
GCC GAT GGT TTG ACC AAT GCC CAG GTA GTC CTA ATT GCT GTT TTC TCC Ala Asp Gly Leu Thr Asn Ala Gln Val Val Leu Ile Ala Val Phe Ser 590 595 600	1947
GTT GCG ATG CCT TTG GTG GCG GTT ATT GCG GCG TGC GTG GTC TTC TGC Val Ala Met Pro Leu Val Ala Val Ile Ala Ala Cys Val Val Phe Cys 605 610 615	1995
ATG AAG CGC AAG CGT AAG CGT GCT CAG GAA AAG GAC GAC GCG GAG GCC Met Lys Arg Lys Arg Lys Arg Ala Gln Glu Lys Asp Asp Ala Glu Ala 620 625 630	2043
AGG AAG CAG AAC GAA CAG AAT GCG GTG GCC ACA ATG CAT CAC AAT GGC Arg Lys Gln Asn Glu Gln Asn Ala Val Ala Thr Met His His Asn Gly 635 640 645 650	2091
AGT GGG GTG GGT GTA GCT TTG GCT TCA GCC TCT CTG GGC GGC AAA ACT Ser Gly Val Gly Val Ala Leu Ala Ser Ala Ser Leu Gly Gly Lys Thr 655 660 665	2139
GGC AGC AAC AGC GGT CTC ACC TTC GAT GGC GGC AAC CCG AAT ATC ATC Gly Ser Asn Ser Gly Leu Thr Phe Asp Gly Gly Asn Pro Asn Ile Ile 670 675 680	2187
AAA AAC ACC TGG GAC AAG TCG GTC AAC AAC ATT TGT GCC TCA GCA GCA Lys Asn Thr Trp Asp Lys Ser Val Asn Asn Ile Cys Ala Ser Ala Ala 685 690 695	2235

FIG. 1 E

6/68

GCA GCG GCG GCG GCG GCA GCA GCG GCG GAC GAG TGT CTC ATG TAC GGC	2283
Ala Ala Ala Ala Ala Ala Ala Ala Ala Asp Glu Cys Leu Met Tyr Gly	
700 705 710	
GGA TAT GTG GCC TCG GTG GCG GAT AAC AAC AAT GCC AAC TCA GAC TTT	2331
Gly Tyr Val Ala Ser Val Ala Asp Asn Asn Asn Ala Asn Ser Asp Phe	
715 720 725 730	
TGT GTG GCT CCG CTA CAA AGA GCC AAG TCG CAA AAG CAA CTC AAC ACC	2379
Cys Val Ala Pro Leu Gln Arg Ala Lys Ser Gln Lys Gln Leu Asn Thr	
735 740 745	
GAT CCC ACG CTC ATG CAC CGC GGT TCG CCG GCA GGC AGC TCA GCC AAG	2427
Asp Pro Thr Leu Met His Arg Gly Ser Pro Ala Gly Ser Ser Ala Lys	
750 755 760	
GGA GCG TCT GGC GGA GGA CCG GGA GCG GCG GAG GGC AAG AGG ATC TCT	2475
Gly Ala Ser Gly Gly Gly Pro Gly Ala Ala Glu Gly Lys Arg Ile Ser	
765 770 775	
GTT TTA GGC GAG GGT TCC TAC TGT AGC CAG CGT TGG CCC TCG TTG GCG	2523
Val Leu Gly Glu Gly Ser Tyr Cys Ser Gln Arg Trp Pro Ser Leu Ala	
780 785 790	
GCG GCG GGA GTG GCC GGA GCC TGT TCA TCC CAG CTA ATG GCT GCA GCT	2571
Ala Ala Gly Val Ala Gly Ala Cys Ser Ser Gln Leu Met Ala Ala Ala	
795 800 805 810	
TCG GCA GCG GGC AGC GGA GCG GGG ACG GCG CAA CAG CAG CGA TCC GTG	2619
Ser Ala Ala Gly Ser Gly Ala Gly Thr Ala Gln Gln Gln Arg Ser Val	
815 820 825	
GTC TGC GGC ACT CCG CAT ATG TAACTCCAAA AATCCGGAAG GGCTCCTGGT	2670
Val Cys Gly Thr Pro His Met	
830	
AAATCCGGAG AAATCCGCAT GGAGGAGCTG ACAGCACATA CACAAAGAAA AGACTGGGTT	2730
GGGTTCAAAA TGTGAGAGAG ACGCCAAAAT GTTGTGTGTG ATTGAAGCAG TTTAGTCGTC	2790
ACGAAAAATG AAAAATCTGT AACAGGCATA ACTCGTAAAC TCCCTAAAAA ATTTGTATAG	2850
TAATTAGCAA AGCTGTGACC CAGCCGTTTC GATCCCGAAT TC	2892

FIG.1F

SUBSTITUTE SHEET (RULE 26)

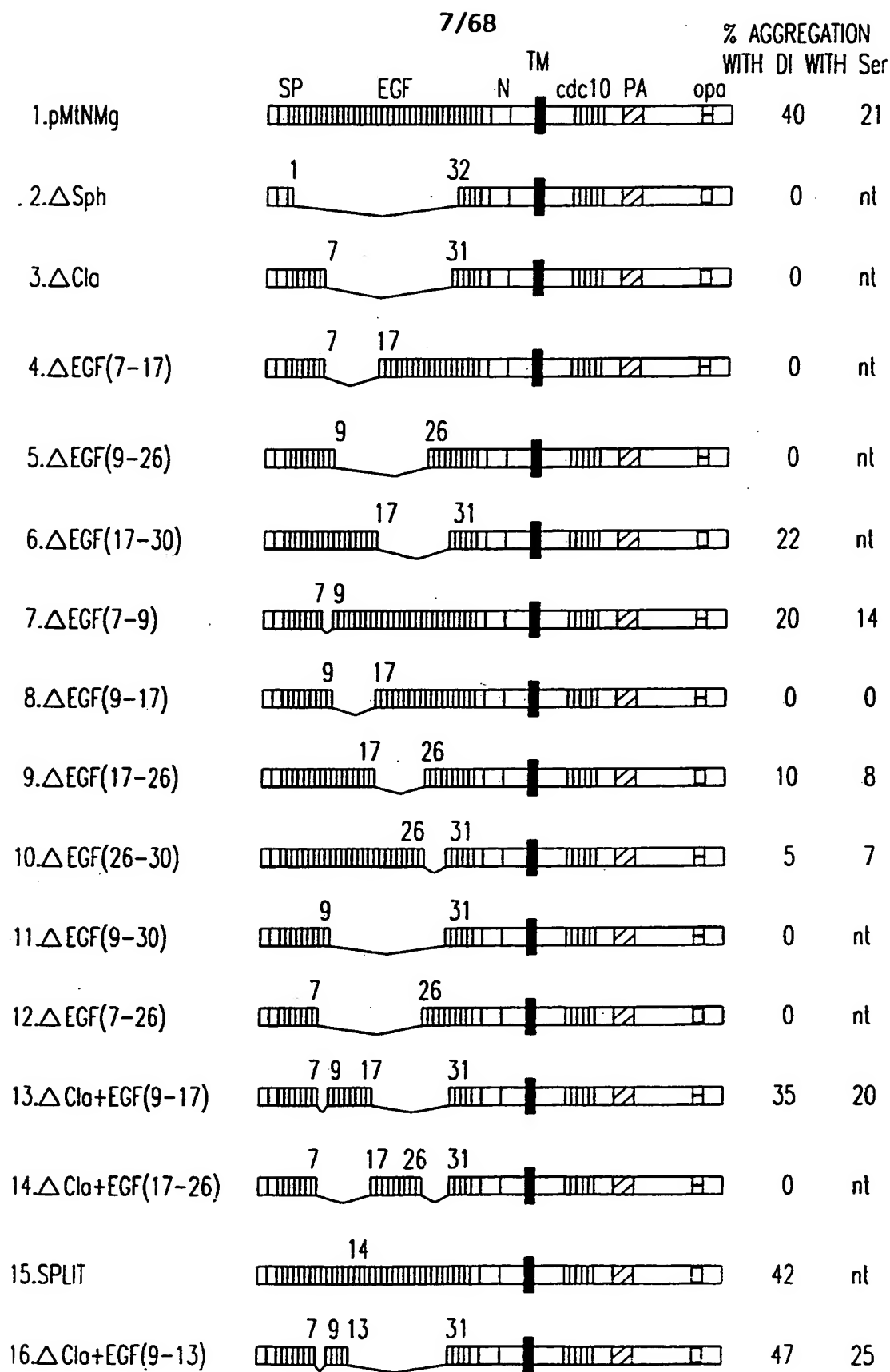


FIG.2A

8/68

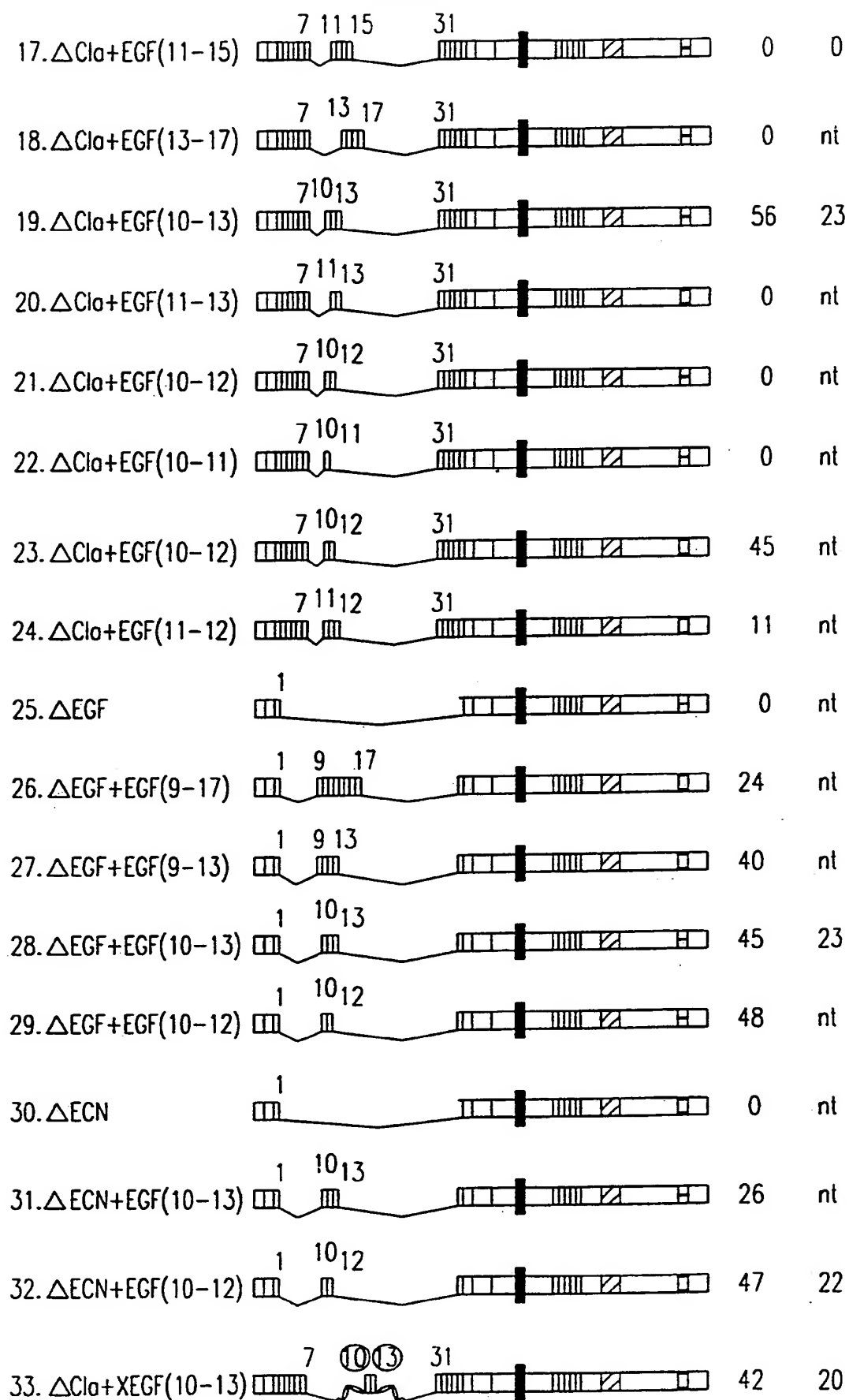
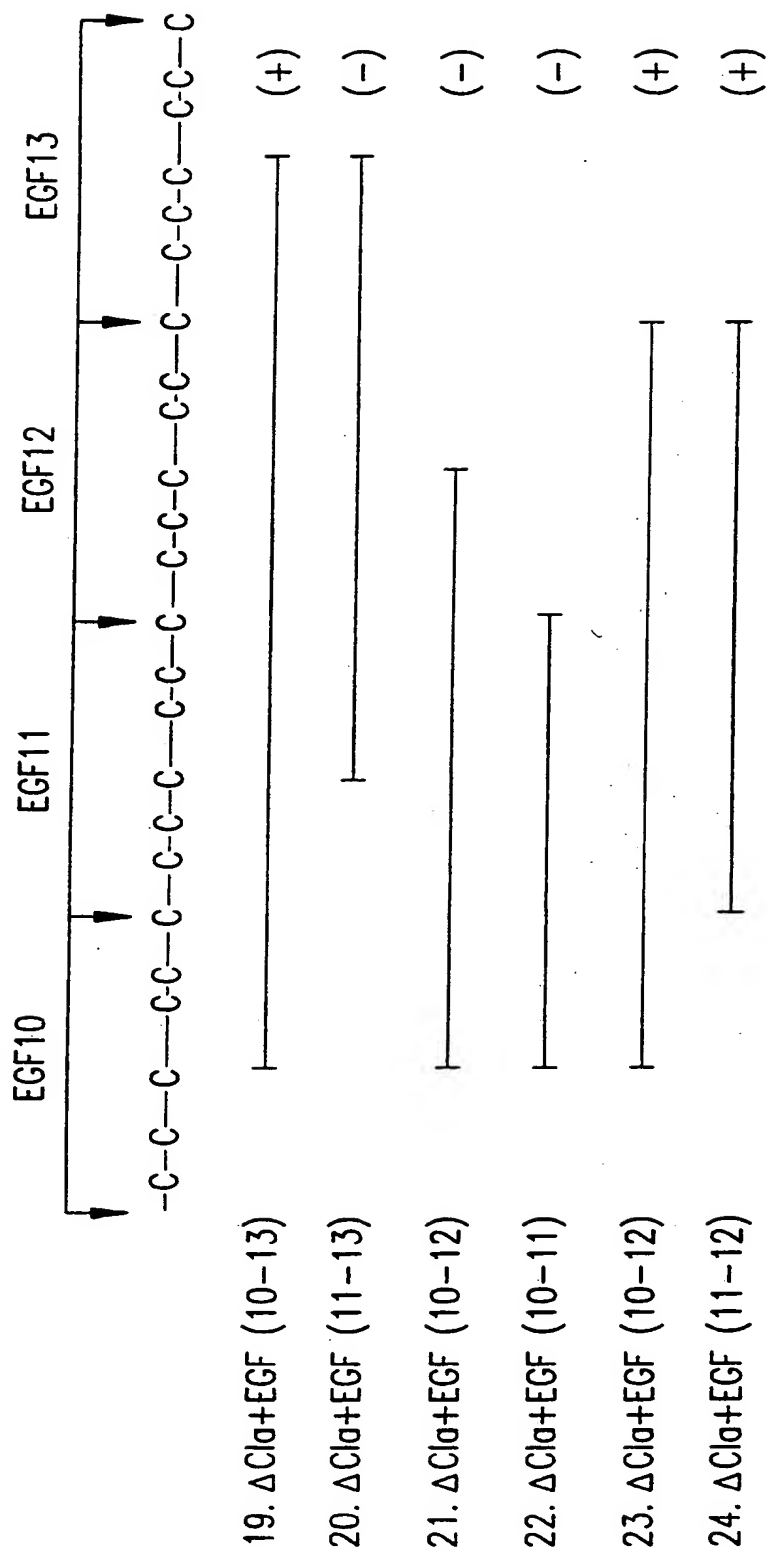


FIG.2B

9/68



**FIG. 3**

10/68

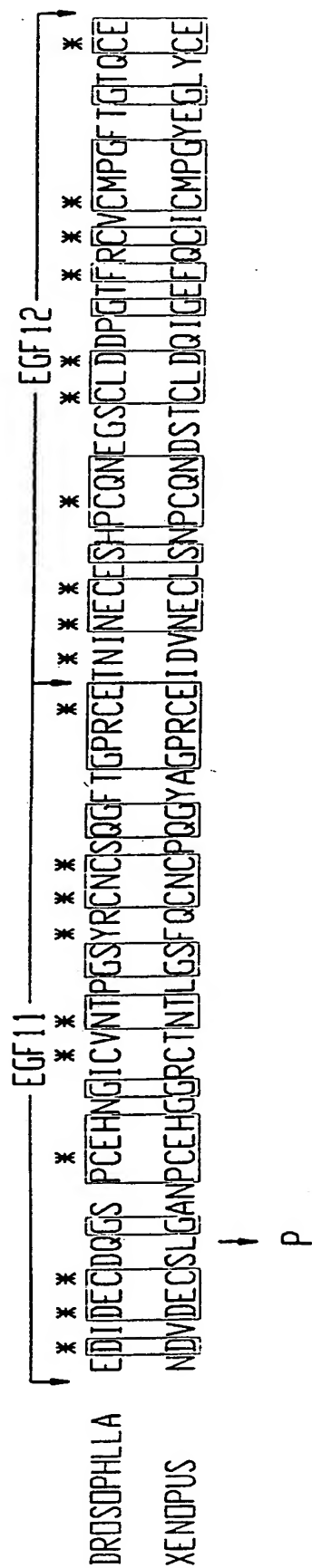


FIG.4



11/68

## FIG. 5A

1 CCGAGTCGAGCGCGGTGCTTCGAGCGGTGATGAGCCCTTTTCGTCAACGCTAAAGATC  
 121 AAGCACATAAGGTCCATATAAATAATAATAATTGTTGTTGATTAACAACATTAT  
 241 GGCCGTTATTCAGCTATCCAGAGCAAGTGTAGTGTGGCAAAATAGAAACAACAAGGCA  
 361 CAATCCAGAGTGAATCCGAACAAACTCCATCTAGATCGCCCAACCAGCATCACGCTCGCA  
  
 481 TCGTCGTTGGAGTCAACAAATAGAAATCAGCAGAGCCTGGGAATGTCCAAGAGACGGCG  
 SerSerLeuGluSerThrIleGluSerAlaAspSerLeuGlyMetSerLysLysThrAla  
  
 601 CCGGATTGTCGATCATTAAGTCTGCCTCGCAACTTAATGCTTTAATTTAAATACGTTA  
 ArgAspCysArgSerLeuLysSerAlaCysAsnLeuIleAlaLeuIleLeuIleLeuLeu  
 -----  
 721 AACAGCCATCTACICACGGCTATTGCTGGGGCATGCCAGCGGAACCTAGGGCCACCAG  
 AsnSerHisLeuLeuAsnGlyTyrCysCysGlyMetProAlaGluLeuArgAlaThrLys  
  
 841 ACCGAGCAGGGTCCAGCATATCCACGGGCTGTTGCTTTGGCAAGCCACCACCAGATA  
 ThrGluGlnGlyAlaSerIleSerThrGlyCysSerPheGlyAsnAlaThrThrLysIle  
 -----  
 961 ACGTTTCGTTGGACGAAGTCGTTACGCTGATCTGCAGCGGTGGATATGTACACACA  
 ThrPheArgTrpThrLysSerPheThrLeuIleLeuGlnAlaLeuAspMetTyrAsnThr  
 -----  
 1081 TCGCCGGAGTGGAGACCCCTGGACCACATCGGGCGGAACGGCGGATCACCTACCGTGT  
 SerProGluTrpLysThrLeuAspHisIleGlyArgAsnAlaArgIleThrTyrArgVal  
 -----  
 1201 GACGATCAGTTCGGTCAGTACGCCTCGGGCTCCGAGGGTCAGAGCTTCGCTGAATGGC  
 AspAspGlnPheGlyHisTyrAlaCysGlySerGluGlyGlnLysLeuCysLeuAsnGly

12/68

TACAAAACATCAGCGCCTATCAAGTGGAGTGTCAGTGTGAACAAACAAAACGAGAG  
 CCAAAACAAACCAAAACAAAGGCAAGTGGAGAAAATGATACAGCATCCAGAGTAC  
 CCAAAATCTGCATACATGGGCTAATTAGGCTGCCAGCGAATTTACATTTGTGTGGTGC  
 AACGCCCCCAGAAATGTACAAAATGTTAGGAACATTTTCGGCGAAACACGCTACGTGG  
 MetPheArgLysHisPheArgArgLysProAlaThrSer 13  
 ACAAAGGAGCGTCCGAGGCATCGGTACCCAAAATCCGACCCCTGCCATCGACGATC  
 ThrLysArgGlnArgProArgHisArgValProLysIleAlaThrLeuProSerThrIle 53  
 GTCCATAAGATATCCGCAGCTGGTAACCTCGAGCTGGAAAATATTAGAAATCTCAATACC  
 ValHisLysIleSerAlaAlaGlyAsnPheGluLeuGluIleSerAsnThr 93  
 -----  
 ACCGATAGGCTGCTCGCCATGCACGACGGCATTCGGGCTGTGCCTGAAGGAGTACCAGACC  
 ThrIleGlyCysSerProCysThrThrAlaPheArgLeuCysLeuLysGluTyrGlnThr 133  
 -----  
 CTGGGTGGCTCCAGCTTGTGCTCAGCGATCCGGGTGTGGAGCCATTGTGCTGCCCTTT  
 LeuGlyGlySerSerPheValLeuSerAspProGlyValGlyAlaIleValLeuProPhe 173  
 -----  
 TCCATCCAGATCGGGAGAGGTAAATTGAGGAACATCATCTCGGGCGGTGATACGCCG  
 SerTyrProAspAlaGluArgLeuIleGluGluThrSerTyrSerGlyValIleLeuPro 213  
 -----  
 CCGGTGCAATGCGCCGTACCTACTACAAACAGACCTGCACGACCTTGTGCCGTCCGGCG  
 ArgValGlnCysAlaValThrTyrTyrAsnThrThrCysThrThrPheCysArgProArg 253  
 -----  
 TGGCAGGGCGTCAACITCGGAGGAGGCCATATGCAAGCGGGGTGGACCCCGTCCACGGC  
 TrpGlnGlyValAsnCysGluGluAlaIleCysLysAlaGlyCysAspProValHisGly 293  
 -----

FIG. 5B

13/68

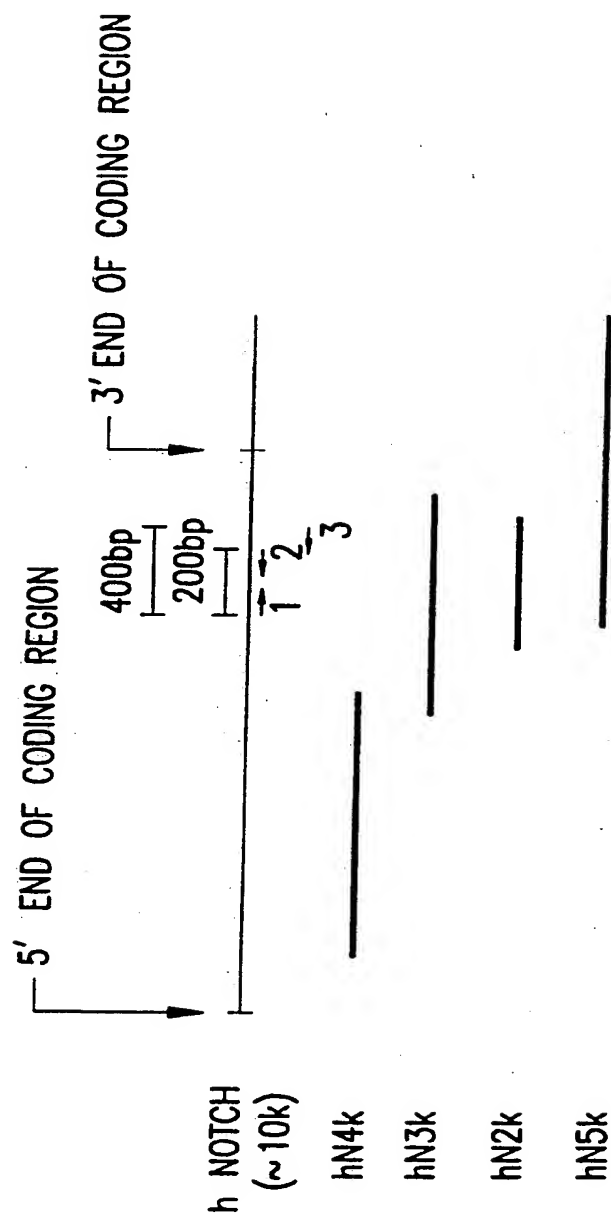
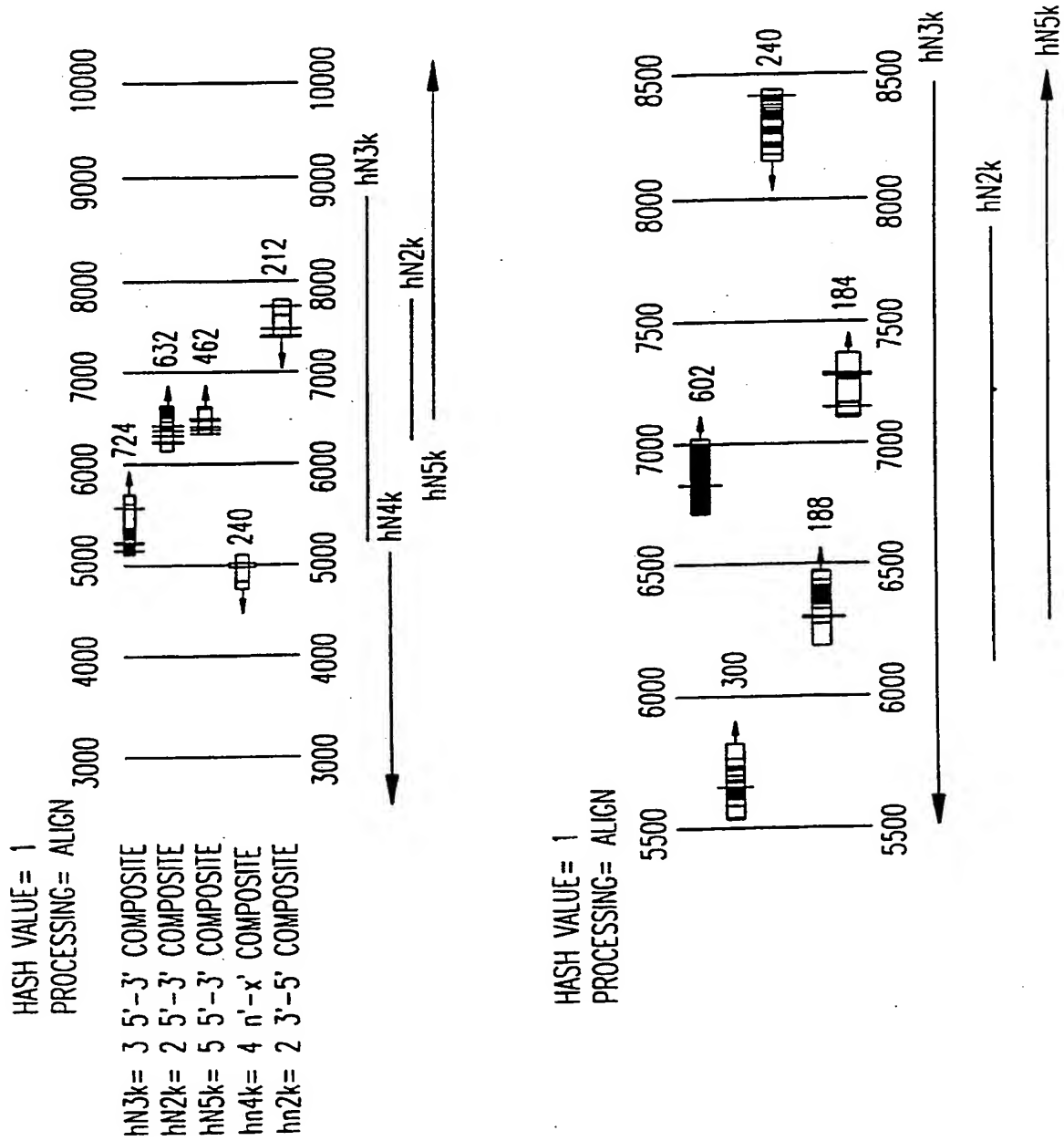


FIG. 6

14/68

FIG. 7



15/68

1 GAATTCGGCT GGGAGAATGG TCTGAGCTAC CTGCCCCTCC TGCTGGGGCA TCAATGGCAA  
61 GTGGGGAAAG CCACACTGGG CAAACGGGCC AGGCCATTTC TGAATGTGG TACATGGTGG  
121 GCAGGGGGCC CGCAACAGCT GGAGGGCAGG TGGACTGAGG CTGGGGATCC CCCGCTGGT  
181 GGGCAATACT GCCTTTACCC ATGAGCTGGA AAGTCACAAT GGGGGGCAAG GGCTCCCGAG  
241 GGTGGTTATG TGCTTCCTTC AGGTGGC

## FIG.8A

1 GAATTCCTTC CATTATACGT GACTTTTCTG AACTGTAGC CACCCTAGTG TCTCTAACTC  
61 CCTCTGGAGT TTGTCAGCTT TGGTCTTTTC AAAGAGCAGG CTCTCTTCAA GTCCTTAAT  
121 GCGGGCATGC TCCAGTTTGG TCTGCGTCTC AAGATCACCT TTGGTAATTG ATTCTTCTTC  
181 AACCCGGAAC TGAAGGCTGG CTCTCACCCT CTAGGCAGAG CAGGAATTCC GAGGTGGATG  
241 TGTTAGATGT GAATGTCCGT GGGCCAGATG GCTGCACCCC ATTGATGTTG GCTTCTCTCC  
301 GAGGAGGCAG CTCAGATTTG AGTGATGAAG ATGAAGATGC AGAGGACTGT TCTGCTAACA  
361 TCATCACAGA CTTGGTCTAC CAGGGTGCCA GCCTCCAGAC CAGACAGACC GGACTGGTGA  
421 GATGGCCCTG CACCTTGCAG CCCGCTACTC ACGGGCTGAT GCTGCCAAGC GTCTCCTGGA  
481 TGCAGGTGCA GATGCCAATG CCCAGGACAA CATGGGCCGC TGTCCACTCC ATGCTGCAGT  
541 GGCACGTGAT GCCAAGGTGT ATTCAGATCT GTTA

## FIG.8B

1 TCCAGATTCT GATTCGCAAC CGAGTAACTG ATCTAGATGC CAGGATGAAT GATGGTACTA  
61 CACCCCTGAT CCTGGCTGCC CGCCTGGCTG TGGAGGGAAT GGTGGCAGAA CTGATCAACT  
121 GCCAAGCGGA TGTGAATGCA GTGGATGACC ATGGAAAATC TGCTCTTAC TGGGCAGCTG  
181 CTGTCAATAA TGTGGAGGCA ACTCTTTTGT TGTGAAAAA TGGGGCCAAC CGAGACATGC  
241 AGGACAACAA GGAAGAGACA CCTCTGTTTC TTGCTGCCCCG GGAGGAGCTA TAAGC

## FIG.8C

16/68

1 GAATTCATT CAGGAGGAAA GGGTGGGAG AGAAGCAGGC ACCCACTTTC CCGTGGCTGG  
61 ACTCGTTCCC AGGTGGCTCC ACCGGCAGCT GTGACCGCCG CAGGTGGGGG CGGAGTGCCA  
121 TTCAGAAAAT TCCAGAAAAG CCCTACCCCA ACTCGGACGG CAACGTCACA CCCGTGGGTA  
181 GCAACTGGCA CACAAACAGC CAGCGTGTCT GGGGCACGGG GGGATGGCAC CCCCTGCAGG  
241 CAGAGCTG

## FIG.9A

1 CTAAAGGGAA CAAAAGCNGG AGCTCCACCG CGGGCGGCNC NGCTCTAGAA CTAGTGGANN  
61 NCCCGGGCTG CAGGAATTCC GCGGGACTGG GCTCGGGCTC AGAGCGGGCG TGTGGAAGAG  
121 ATTCTAGACC GGGAGAACAA GCGAATGGCT GACAGCTGGC CTCCAAAGTC ACCAGGCTCA  
181 AATCGCTCGC CCTGGACATC GAGGGATGCA GAGGATCAGA ACCGGTACCT GGATGGCATG  
241 ACTCGGATTT ACAAGCATGA CCAGCCTGCT TACAGGGAGC GTGANNTTTT CACATGCAGT  
301 CGACAGACAC GAGCTCTATG CAT

## FIG.9B

17/68

```

10      *      *      *      *      *      *      *      *      *      *
      TGC CAG GAG GAC GCG GGC AAC AAG GTC TGC AGC CTG CAG TGC AAC AAC
      C   Q   E   D   A   G   N   K   V   C   S   L   Q   C   N   N>

50      *      *      *      *      *      *      *      *      *      *
      CAC GCG TGC GGC TGG GAC GGC GGT GAC TGC TCC CTC AAC TTC AAT GAC
      H   A   C   G   W   D   G   G   D   C   S   L   N   F   N   D>

100     *      *      *      *      *      *      *      *      *      *
      CCC TGG AAG AAC TGC ACG CAG TCT CTG CAG TGC TGG AAG TAC TTC AGT
      P   W   K   N   C   T   Q   S   L   Q   C   W   K   Y   F   S>

150     *      *      *      *      *      *      *      *      *      *
      GAC GGC CAC TGT GAC AGC CAG TGC AAC TCA GCC GGC TGC CTC TTC GAC
      D   G   H   C   D   S   Q   C   N   S   A   G   C   L   F   D>

```

FIG. 10A

18/68

```

200      *      *      *      *      *      *      *      *      *      *
      GGC TTT GAC TGC CAG CGT GCG GAA GGC CAG TGC AAC CCC CTG TAC GAC
      G F D C Q R A E G Q C N P L Y D> -
210      *      *      *      *      *      *      *      *      *      *
220      *      *      *      *      *      *      *      *      *      *
230      *      *      *      *      *      *      *      *      *      *
240      *      *      *      *      *      *      *      *      *      *

250      *      *      *      *      *      *      *      *      *      *
      CAG TAC TGC AAG GAC CAC TTC AGC GAC GGC CAC TGC GAC CAG GGC TGC
      Q Y C K D H F S D G H C D Q G C>
260      *      *      *      *      *      *      *      *      *      *
270      *      *      *      *      *      *      *      *      *      *
280      *      *      *      *      *      *      *      *      *      *

290      *      *      *      *      *      *      *      *      *      *
      AAC AGC GCG GAG TGC GAG TGG GAC GGC CTG GAC TGT GCG GAG CAT GTA
      N S A E C E W D G L D C A E H V>
300      *      *      *      *      *      *      *      *      *      *
310      *      *      *      *      *      *      *      *      *      *
320      *      *      *      *      *      *      *      *      *      *
330      *      *      *      *      *      *      *      *      *      *

340      *      *      *      *      *      *      *      *      *      *
      CCC GAG AGG CTG GCG GCC GGC ACC CTG GTG GTG GTG CTG ATG CCG
      P E R L A A G T L V V V V L M P>
350      *      *      *      *      *      *      *      *      *      *
360      *      *      *      *      *      *      *      *      *      *
370      *      *      *      *      *      *      *      *      *      *
380      *      *      *      *      *      *      *      *      *      *

```

FIG.10B



19/68

```

390 *      *      *      *      *      *      *      *      *      *
    CCG GAG CAG CTG CGC AAC AGC TCC TTC CAC TTC CTG CGG GAG CTC AGC
    P   E   Q   L   R   N   S   S   F   H   F   L   R   E   L   S>

440 *      *      *      *      *      *      *      *      *      *
    CGC GTG CTG CAC ACC AAC GTG GTC TTC AAG CGT GAC GCA CAC GGC CAG
    R   V   L   H   T   N   V   V   F   K   R   D   A   H   G   Q>

490 *      *      *      *      *      *      *      *      *      *
    CAG ATG ATC TTC CCC TAC TAC GGC CGC GAG GAG GAG CTG CGC AAG CAC
    Q   M   I   F   P   Y   Y   G   R   E   E   L   R   K   H>

530 *      *      *      *      *      *      *      *      *      *
    CCC ATC AAG CGT GCC GCC GAG GGC TGG GCC GCA CCT GAC GCC CTG CTG
    P   I   K   R   A   A   E   G   W   A   A   P   D   A   L   L>

```

FIG. 10C

20/68

```

580      *      *      *      *      *      *      *      *      *      *
      GGC CAG GTG AAG GCC TCG CTG CTC CCT GGT GGC AGC GAG GGT GGG CGG
      G  Q  V  K  A  S  L  L  P  G  G  S  E  G  G  R>

630      *      *      *      *      *      *      *      *      *      *
      CGG CGG AGG GAG CTG GAC CCC ATG GAC GTC CGC GGC TCC ATC GTC TAC
      R  R  E  L  D  P  M  D  V  R  G  S  I  V  Y>

680      *      *      *      *      *      *      *      *      *      *
      CTG GAG ATT GAC AAC CCG CAG TGT GTG CAG GCC TCC TCG CAG TGC TTC
      L  E  I  D  N  R  Q  C  V  Q  A  S  S  Q  C  F>

730      *      *      *      *      *      *      *      *      *      *
      CAG AGT GCC ACC GAC GTG GCC GCA TTC CTG GGA GCG CTC GCC TCG CTG
      Q  S  A  T  D  V  A  A  F  L  G  A  L  A  S  L>

```

FIG. 10D

21/68

```

770      *      *      *      *      *      *      *      *      *      *
      GGC AGC CTC AAC ATC CCC TAC AAG ATC GAG GCC GTG CAG AGT GAG ACC
      G   S   L   N   I   P   Y   K   I   E   A   V   Q   S   E   T>~

      820      *      *      *      *      *      *      *      *      *      *
      GTG GAG CCG CCC CCG CCG CAG CTG CAC TTC ATG TAC GTG GCG GCG
      V   E   P   P   P   A   Q   L   H   F   M   Y   V   A   A>

      870      *      *      *      *      *      *      *      *      *      *
      GCC GCC TTT GTG CTT CTG TTC TTC GTG GGC TGC GGG GTG CTG CTG TCC
      A   A   F   V   L   L   F   F   V   G   C   G   V   L   L   S>

      920      *      *      *      *      *      *      *      *      *      *
      CGC AAG CGC CGG CGG CAG CAT GGC CAG CTC TGG TTC CCT GAG GGC TTC
      R   K   R   R   R   Q   H   G   Q   L   W   F   P   E   G   E>

      830      *      *      *      *      *      *      *      *      *      *
      840      *      *      *      *      *      *      *      *      *      *
      850      *      *      *      *      *      *      *      *      *      *
      860      *      *      *      *      *      *      *      *      *      *
      880      *      *      *      *      *      *      *      *      *      *
      890      *      *      *      *      *      *      *      *      *      *
      900      *      *      *      *      *      *      *      *      *      *
      910      *      *      *      *      *      *      *      *      *      *
      930      *      *      *      *      *      *      *      *      *      *
      940      *      *      *      *      *      *      *      *      *      *
      950      *      *      *      *      *      *      *      *      *      *
      960      *      *      *      *      *      *      *      *      *      *

```

FIG.10E

22/68

```

          970      980      990      1000
*   *   *   *   *   *   *   *
AAA GTG TCT GAG GCC AGC AAG AAG AAG CGG CGG GAG CCC CTC GGC GAG
K   V   S   E   A   S   K   K   K   R   R   E   P   L   G   E>

1010      1020      1030      1040      1050
*   *   *   *   *   *   *   *
GAC TCC GTG GGC CTC AAG CCC CTG AAG AAC GCT TCA GAC GGT GCC CTC
D   S   V   G   L   K   P   L   K   N   A   S   D   G   A   L>

1060      1070      1080      1090      1100
*   *   *   *   *   *   *   *
ATG GAC GAC AAC CAG AAT GAG TGG GGG GAC GAG GAC CTG GAG ACC AAG
M   D   D   N   Q   N   E   W   G   D   E   D   L   E   T   K>

1110      1120      1130      1140      1150
*   *   *   *   *   *   *   *
AAG TTC CGG TTC GAG GAG CCC GTG GTT CTG CCT GAC CTG GAC GAC CAG
K   F   R   F   E   E   P   V   V   L   P   D   L   D   D   Q>

```

FIG.10F

23/68

```

1160      *      1170      *      1180      *      1190      *      1200      *
ACA GAC CAC CGG CAG TGG ACT CAG CAG CAC CTG GAT GCC GCT GAC CTG
T   D   H   R   Q   W   T   Q   Q   H   L   D   A   A   D   L>

1210      *      1220      *      1230      *      1240      *
CGC ATG TCT GCC ATG GCC CCC ACA CCG CCC CAG GGT GAG GTT GAC GCC
R   M   S   A   M   A   P   T   P   P   Q   G   E   V   D   A>

1250      *      1260      *      1270      *      1280      *      1290      *
GAC TGC ATG GAC GTC AAT GTC CGC GGC CCT GAT GGC TTC ACC CCG CTC
D   C   M   D   V   N   V   R   G   G   P   D   G   F   T   P   L>

1300      *      1310      *      1320      *      1330      *      1340      *
ATG ATC GCC TCC TGC AGC GGC GGC GGC CTG GAG ACG GGC AAC AGC GAG
M   I   A   S   C   S   G   G   G   L   E   T   G   N   S   E>

```

FIG.10G

24/68

```

1350      *      *      *      *      *      *      *      *      *      *
      GAA GAG GAG GAC GCG CCG GCC GTC ATC TCC GAC TTC ATC TAC CAG GGC
      E   E   D   A   P   A   V   I   S   D   F   I   Y   Q   G>

1400      *      *      *      *      *      *      *      *      *      *
      GCC AGC CTG CAC AAC CAG ACA GAC CGC ACG GGC GAG ACC GCC TTG CAC
      A   S   L   H   N   Q   T   D   R   T   G   E   T   A   L   HD>

1450      *      *      *      *      *      *      *      *      *      *
      CTG GCC GCC CGC TAC TCA CGC TCT GAT GCC GCC AAG CGC CTG CTG GAG
      L   A   A   R   Y   S   R   S   D   A   A   K   R   L   L   E>

1490      *      *      *      *      *      *      *      *      *      *
      GCC AGC GCA GAT GCC AAC ATC CAG GAC AAC ATG GGC CGC ACC CCG CTG
      A   S   A   D   A   N   I   Q   D   N   M   G   R   T   P   LD>

```

FIG. 10H

25/68

```

1540      *      1550      *      1560      *      1570      *      1580      *
CAT GCG GCT GTG TCT GCC GAC GCA CAA GGT GTC TTC CAG ATC CTG ATC
H   A   A   V   S   A   D   A   Q   G   V   F   Q   I   L   I>

1590      *      1600      *      1610      *      1620      *      1630      *
CGG AAC CGA GCC ACA GAC CTG GAT GCC CGC ATG CAT GAT GGC ACG ACG
R   N   R   A   T   D   L   D   A   R   M   H   D   G   T   T>

1640      *      1650      *      1660      *      1670      *      1680      *
CCA CTG ATC CTG GCT GCC CGC CTG GCC GTG GAG GGC ATG CTG GAG GAC
P   L   I   L   A   A   R   L   A   V   E   G   M   L   E   D>

1690      *      1700      *      1710      *      1720      *
CTC ATC AAC TCA CAC GCC GAC GTC AAC GCC GTA GAT GAC CTG GGC AAG
L   I   N   S   H   A   D   V   N   A   V   D   D   L   G   K>

1730      *      1740      *      1750      *      1760      *      1770      *
TCC GCC CTG CAC TGG GCC GCC GGC GTG AAC AAT GTG GAT GCC GCA GTT
S   A   L   H   W   A   A   A   V   N   N   V   D   A   A   V>

```

FIG. 10I

26/68

1780	1790	1800	1810	1820
* GTG CTC CTG AAG AAC GGG GCT AAC AAA GAT ATG CAG AAC AAC AGG GAG	* V L L K N G A N K D M Q N R E>			
1830	1840	1850	1860	1870
* GAG ACA CCC CTG TTT CTG GCC GCC CGG GAG GGC AGC TAC GAG ACC GCC	* E T P L F L A A R E G S Y E T A>			
1880	1890	1900	1910	1920
* AAG GTG CTG CTG GAC CAC TTT GCC AAC CGG GAC ATC ACG GAT CAT ATG	* K V L L D H F A N R D I T D H M>			
1930	1940	1950	1960	
* GAC CGC CTG CCG CGC GAC ATC GCA CAG GAG CGC ATG CAT CAC GAC ATC	* D R L P R D I A Q E R M H H D I>			

FIG.10J



27/68

1970	1980	1990	2000	2010	
* GTG AGG CTG GAC GAG TAC AAC CTG GTG CGC AGC CCG CAG CTG CAC	* GTG AGG CTG GAC GAG TAC AAC CTG GTG CGC AGC CCG CAG CTG CAC	* GTG AGG CTG GAC GAG TAC AAC CTG GTG CGC AGC CCG CAG CTG CAC	* GTG AGG CTG GAC GAG TAC AAC CTG GTG CGC AGC CCG CAG CTG CAC	* GTG AGG CTG GAC GAG TAC AAC CTG GTG CGC AGC CCG CAG CTG CAC	
V R L L D E Y N L V R S P Q L HD	V R L L D E Y N L V R S P Q L HD	V R L L D E Y N L V R S P Q L HD	V R L L D E Y N L V R S P Q L HD	V R L L D E Y N L V R S P Q L HD	
2020	2030	2040	2050	2060	
* GGA GCC CCG CTG GGG GGC ACG CCC ACC CTG TCG CCC CCG CTC TGC TCG	* GGA GCC CCG CTG GGG GGC ACG CCC ACC CTG TCG CCC CCG CTC TGC TCG	* GGA GCC CCG CTG GGG GGC ACG CCC ACC CTG TCG CCC CCG CTC TGC TCG	* GGA GCC CCG CTG GGG GGC ACG CCC ACC CTG TCG CCC CCG CTC TGC TCG	* GGA GCC CCG CTG GGG GGC ACG CCC ACC CTG TCG CCC CCG CTC TGC TCG	
G A P L G G T P T L S P P L C S>	G A P L G G T P T L S P P L C S>	G A P L G G T P T L S P P L C S>	G A P L G G T P T L S P P L C S>	G A P L G G T P T L S P P L C S>	
2070	2080	2090	2100	2110	
* CCC AAC GGC TAC CTG GGC AGC CTC AAG CCC GGC GTG CAG GGC AAG AAG	* CCC AAC GGC TAC CTG GGC AGC CTC AAG CCC GGC GTG CAG GGC AAG AAG	* CCC AAC GGC TAC CTG GGC AGC CTC AAG CCC GGC GTG CAG GGC AAG AAG	* CCC AAC GGC TAC CTG GGC AGC CTC AAG CCC GGC GTG CAG GGC AAG AAG	* CCC AAC GGC TAC CTG GGC AGC CTC AAG CCC GGC GTG CAG GGC AAG AAG	
P N G Y L L S L K P P G V Q G K KD	P N G Y L L S L K P P G V Q G K KD	P N G Y L L S L K P P G V Q G K KD	P N G Y L L S L K P P G V Q G K KD	P N G Y L L S L K P P G V Q G K KD	
2120	2130	2140	2150	2160	
* GTC CGC AAG CCC AGC AGC AAA GGC CTG GCC TGT GGA AGC AAG GAG GCC	* GTC CGC AAG CCC AGC AGC AAA GGC CTG GCC TGT GGA AGC AAG GAG GCC	* GTC CGC AAG CCC AGC AGC AAA GGC CTG GCC TGT GGA AGC AAG GAG GCC	* GTC CGC AAG CCC AGC AGC AAA GGC CTG GCC TGT GGA AGC AAG GAG GCC	* GTC CGC AAG CCC AGC AGC AAA GGC CTG GCC TGT GGA AGC AAG GAG GCC	
V R K P S S K G G L A C G S K E AD	V R K P S S K G G L A C G S K E AD	V R K P S S K G G L A C G S K E AD	V R K P S S K G G L A C G S K E AD	V R K P S S K G G L A C G S K E AD	

FIG. 10K

28/68

2170	2180	2190	2200
* AAG GAC CTC AAG GCA CGG AGG AAG AAG TCC CAG GAT GGC AAG GGC TGC	* K D L K A R R K K S Q D G K G C>		
2210	2220	2230	2240
* CTG CTG GAC AGC TCC GGC ATG CTC TCG CCC GTG GAC TCC CTG GAG TCA	* L D S S G M L S P V D S L E S>		
2260	2270	2280	2290
* CCC CAT GGC TAC CTG TCA GAC GTG GCC TCG CCG CCA CTG CTG CCC TCC	* P H G Y L S D V A S P P L L P S>		
2310	2320	2330	2340
* CCG TTC CAG CAG TCT CCG TCC GTG CCC CTC AAC CAC CTG CCT GGG ATG	* P F Q Q S P S V P L N H L P G MD		

FIG.10L

29/68

2360	2370	2380	2390	2400
* CCC GAC ACC CAC CTG GGC ATC GGG CAC CTG AAC GTG GCG GCC AAG CCC				*
P D T H L G I G H L N V A A K P>				
2410	2420	2430	2440	
* GAG ATG GCG GCG CTG GGT GGC GGC GCG CTG GCC TTT GAG ACT GGC			*	
E M A A L G G G G R L A F E T G>				
2450	2460	2470	2480	2490
* CCA CCT CGT CTC TCC CAC CTG CCT CTG GGC TCT GGC ACC AGC ACC GTC			*	*
P P R L S H L P V A S G T S T V>				
2500	2510	2520	2530	2540
* CTG GGC TCC AGC AGC GGA GGC GGC CTG AAT TTC ACT GTG GGC GGC TCC			*	*
L G S S S G G A L N F T V G G S>				

FIG.10M

30/68

2550	2560	2570	2580	2590
* * *	* * *	* * *	* * *	* * *
ACC AGT TTG AAT GGT CAA TGC GAG TGG CTG TCC CGG CTG CAG AGC GGC				
T S L N G Q C E W L S R L Q S G>				
2600	2610	2620	2630	2640
* * *	* * *	* * *	* * *	* * *
ATG GTG CCG AAC CAA TAC AAC CCT CTG CCG GGG AGT GTG GCA CCA GGC				
M V P N Q Y N P L R G S V A P G>				
2650	2660	2670	2680	
* * *	* * *	* * *	* * *	
CCC CTG AGC ACA CAG GCC CCC TCC CTG CAG CAT GGC ATG GTA GGC CCG				
P L S T Q A P S L Q H G M V G P>				
2690	2700	2710	2720	2730
* * *	* * *	* * *	* * *	* * *
CTG CAC AGT AGC CTT GCT GCC AGC GCC CTG TCC CAG ATG ATG AGC TAC				
L H S S L A A S A L S Q M M S Y>				

FIG. 10N

31/68

2740	2750	2760	2770	2780
* * *	* * *	* * *	* * *	* * *
CAG GGC CTG CCC AGC ACC CGG CTG GCC ACC CAG CCT CAC CTG GTG CAG				
Q G L P S T R L A T Q P H L V Q>				
2790	2800	2810	2820	2830
* * *	* * *	* * *	* * *	* * *
ACC CAG CAG GTG CAG CCA CAA AAC TTA CAG ATG CAG CAG AAC CTG				
T Q Q V Q P Q N L Q M Q Q N L>				
2840	2850	2860	2870	2880
* * *	* * *	* * *	* * *	* * *
CAG CCA GCA AAC ATC CAG CAG CAG CAA AGC AGC CTG CAG CCG CCA CCA CCA				
Q P A N I Q Q Q Q S L Q P P P>				
2890	2900	2910	2920	
* * *	* * *	* * *	* * *	
CCA CCA CAG CCG CAC CTT GGC GTG AGC TCA GCA GCC AGC GGC CAC CTG				
P P Q P H L G V S S A A S G H L>				
2930	2940	2950	2960	2970
* * *	* * *	* * *	* * *	* * *
GGC CCG AGC TTC CTG AGT GGA GAG CCG AGC CAG GCA GAC GTG CAG CCA				
G R S F L S G E P S Q A D V Q P>				

FIG.100

32 / 68

2980	2990	3000	3010	3020
CTG GGC CCC AGC AGC CTG GCG GTG CAC ACT ATT CTG CCC CAG GAG AGC				
L G P S S L A V H T I L P Q E S>				
3030	3040	3050	3060	3070
CCC GCC CTG CCC ACG TCG TCG CTG CCA TCC TCG CTG CCA CCC GTG ACC				
P A L P T S L P S S L V P V T>				
3080	3090	3100	3110	3120
GCA GCC CAG TTC CTG ACG CCC CCC TCG CAG CAC AGC TAC TCC TCG CCT				
A A Q F L T P P S Q H S Y S P>				

FIG.10P

33/68

3130	3140	3150	3160
* GTG GAC AAC ACC CCC AGC CAC CAG CTA CAG GTG CCT GTT CCT GTA ATG	* * * *		
V D N T P S H Q L Q V P V M>			
3170	3180	3190	3200
* * * *	* * * *		
GTA ATG ATC CGA TCT TCG GAT CCT TCT AAA GGC TCA ATT TTG ATC			
V M I R S S D P S K G S I L I>			
3220	3230		
* * *	* *		
GAA GCT CCC GAC TCA TGG			
E A P D S W>			

FIG.10Q

34/68

G GAG GTG GAT GTG TTA GAT GTG AAT GTC CGT GGC CCA GAT GGC TGC	46
Glu Val Asp Val Leu Asp Val Asn Val Arg Gly Pro Asp Gly Cys	
1 5 10 15	
ACC CCA TTG ATG TTG GCT TCT CTC CGA GGA GGC AGC TCA GAT TTG AGT	94
Thr Pro Leu Met Leu Ala Ser Leu Arg Gly Gly Ser Ser Asp Leu Ser	
20 25 50	
GAT GAA GAT GAA GAT GCA GAG GAC TCT TCT GCT AAC ATC ATC ACA GAC	142
Asp Glu Asp Glu Asp Ala Glu Asp Ser Ser Ala Asn Ile Ile Thr Asp	
35 40 45	
TTG GTC TAC CAG GGT GCC AGC CTC CAG GCC CAG ACA GAC CGG ACT GGT	190
Leu Val Tyr Gln Gly Ala Ser Leu Gln Ala Gln Thr Asp Arg Thr Gly	
50 55 60	
GAG ATG GCC CTG CAC CTT GCA GCC CGC TAC TCA CGG GCT GAT GCT GCC	238
Glu Met Ala Leu His Leu Ala Ala Arg Tyr Ser Arg Ala Asp Ala Ala	
65 70 75	
AAG CGT CTC CTG GAT GCA GGT GCA GAT GCC AAT GCC CAG GAC AAC ATG	286
Lys Arg Leu Leu Asp Ala Gly Ala Asp Ala Asn Ala Gln Asp Asn Met	
80 85 90 95	
GGC CGC TGT CCA CTC CAT GCT GCA GTG GCA GCT GAT GCC CAA GGT GTC	334
Gly Arg Cys Pro Leu His Ala Ala Val Ala Ala Asp Ala Gln Gly Val	
100 105 110	
TTC CAG ATT CTG ATT CGC AAC CGA GTA ACT GAT CTA GAT GCC AGG ATG	382
Phe Gln Ile Leu Ile Arg Asn Arg Val Thr Asp Leu Asp Ala Arg Met	
115 120 125	
AAT GAT GGT ACT ACA CCC CTG ATC CTG GCT GCC CGC CTG GCT GTG GAG	430
Asn Asp Gly Thr Thr Pro Leu Ile Leu Ala Ala Arg Leu Ala Val Glu	
130 135 140	

FIG.11A



35/68

GGA ATG GTG GCA GAA CTG ATC AAC TGC CAA GCG GAT GTG AAT GCA GTG	478
Gly Met Val Ala Glu Leu Ile Asn Cys Gln Ala Asp Val Asn Ala Val	
145 150 155	
GAT GAC CAT GGA AAA TCT GCT CTT CAC TGG GCA GCT GCT GTC AAT AAT	526
Asp Asp His Gly Lys Ser Ala Leu His Trp Ala Ala Ala Val Asn Asn	
160 165 170 175	
GTG GAG GCA ACT CTT TTG TTG TTG AAA AAT GGG GCC AAC CGA GAC ATG	574
Val Glu Ala Thr Leu Leu Leu Leu Lys Asn Gly Ala Asn Arg Asp Met	
180 185 190	
CAG GAC AAC AAG GAA GAG ACA CCT CTG TTT CTT GCT GCC CGG GAG GGG	622
Gln Asp Asn Lys Glu Glu Thr Pro Leu Phe Leu Ala Ala Arg Glu Gly	
195 200 205	
AGC TAT GAA GCA GCC AAG ATC CTG TTA GAC CAT TTT GCC AAT CGA GAC	670
Ser Tyr Glu Ala Ala Lys Ile Leu Leu Asp His Phe Ala Asn Arg Asp	
210 215 220	
ATC ACA GAC CAT ATG GAT CGT CTT CCC CGG GAT GTG GCT CGG GAT CGC	718
Ile Thr Asp His Met Asp Arg Leu Pro Arg Asp Val Ala Arg Asp Arg	
225 230 235	
ATG CAC CAT GAC ATT GTG CGC CTT CTG GAT GAA TAC AAT GTG ACC CCA	766
Met His His Asp Ile Val Arg Leu Leu Asp Glu Tyr Asn Val Thr Pro	
240 245 250 255	
AGC CCT CCA GGC ACC GTG TTG ACT TCT GCT CTC TCA CCT GTC ATC TGT	814
Ser Pro Pro Gly Thr Val Leu Thr Ser Ala Leu Ser Pro Val Ile Cys	
260 265 270	
GGG CCC AAC AGA TCT TTC CTC AGC CTG AAG CAC ACC CCA ATG GGC AAG	862
Gly Pro Asn Arg Ser Phe Leu Ser Leu Lys His Thr Pro Met Gly Lys	
275 280 285	

FIG.11B

36/68

AAG TCT AGA CGG CCC AGT GCC AAG AGT ACC ATG CCT ACT AGC CTC CCT	910
Lys Ser Arg Arg Pro Ser Ala Lys Ser Thr Met Pro Thr Ser Leu Pro	
290 295 300	
AAC CTT GCC AAG GAG GCA AAG GAT GCC AAG GGT AGT AGG AGG AAG AAG	958
Asn Leu Ala Lys Glu Ala Lys Asp Ala Lys Gly Ser Arg Arg Lys Lys	
305 310 315	
TCT CTG AGT GAG AAG GTC CAA CTG TCT GAG AGT TCA GTA ACT TTA TCC	1006
Ser Leu Ser Glu Lys Val Gln Leu Ser Glu Ser Ser Val Thr Leu Ser	
320 325 330 335	
CCT GTT GAT TCC CTA GAA TCT CCT CAC ACG TAT GTT TCC GAC ACC ACA	1054
Pro Val Asp Ser Leu Glu Ser Pro His Thr Tyr Val Ser Asp Thr Thr	
340 345 350	
TCC TCT CCA ATG ATT ACA TCC CCT GGG ATC TTA CAG GCC TCA CCC AAC	1102
Ser Ser Pro Met Ile Thr Ser Pro Gly Ile Leu Gln Ala Ser Pro Asn	
355 360 365	
CCT ATG TTG GCC ACT GCC GCC CCT CCT GCC CCA GTC CAT GCC CAG CAT	1150
Pro Met Leu Ala Thr Ala Ala Pro Pro Ala Pro Val His Ala Gln His	
370 375 380	
GCA CTA TCT TTT TCT AAC CTT CAT GAA ATG CAG CCT TTG GCA CAT GGG	1198
Ala Leu Ser Phe Ser Asn Leu His Glu Met Gln Pro Leu Ala His Gly	
385 390 395	
GCC AGC ACT GTG CTT CCC TCA GTG AGC CAG TTG CTA TCC CAC CAC CAC	1246
Ala Ser Thr Val Leu Pro Ser Val Ser Gln Leu Leu Ser His His His	
400 405 410 415	
ATT GTG TCT CCA GGC AGT GGC AGT GCT GGA AGC TTG AGT AGG CTC CAT	1294
Ile Val Ser Pro Gly Ser Gly Ser Ala Gly Ser Leu Ser Arg Leu His	
420 425 430	
CCA GTC CCA GTC CCA GCA GAT TGG ATG AAC CGC ATG GAG GTG AAT GAG	1342
Pro Val Pro Val Pro Ala Asp Trp Met Asn Arg Met Glu Val Asn Glu	
435 440 445	

FIG.11C

37/68

ACC CAG TAC AAT GAG ATG TTT GGT ATG GTC CTG GCT CCA GCT GAG GGC	1390
Thr Gln Tyr Asn Glu Met Phe Gly Met Val Leu Ala Pro Ala Glu Gly	
450 455 460	
ACC CAT CCT GGC ATA GCT CCC CAG AGC AGG CCA CCT GAA GGG AAG CAC	1438
Thr His Pro Gly Ile Ala Pro Gln Ser Arg Pro Pro Glu Gly Lys His	
465 470 475	
ATA ACC ACC CCT CGG GAG CCC TTG CCC CCC ATT GTG ACT TTC CAG CTC	1486
Ile Thr Thr Pro Arg Glu Pro Leu Pro Pro Ile Val Thr Phe Gln Leu	
480 485 490 495	
ATC CCT AAA GGC AGT ATT GCC CAA CCA GCG GGG GCT CCC CAG CCT CAG	1534
Ile Pro Lys Gly Ser Ile Ala Gln Pro Ala Gly Ala Pro Gln Pro Gln	
500 505 510	
TCC ACC TGC CCT CCA GCT GTT GCG GGC CCC CTG CCC ACC ATG TAC CAG	1582
Ser Thr Cys Pro Pro Ala Val Ala Gly Pro Leu Pro Thr Met Tyr Gln	
515 520 525	
ATT CCA GAA ATG GCC CGT TTG CCC AGT GTG GCT TTC CCC ACT GCC ATG	1630
Ile Pro Glu Met Ala Arg Leu Pro Ser Val Ala Phe Pro Thr Ala Met	
530 535 540	
ATG CCC CAG CAG GAC GGG CAG GTA GCT CAG ACC ATT CTC CCA GCC TAT	1678
Met Pro Gln Gln Asp Gly Gln Val Ala Gln Thr Ile Leu Pro Ala Tyr	
545 550 555	
CAT CCT TTC CCA GCC TCT GTG GGC AAG TAC CCC ACA CCC CCT TCA CAG	1726
His Pro Phe Pro Ala Ser Val Gly Lys Tyr Pro Thr Pro Pro Ser Gln	
560 565 570 575	
CAC AGT TAT GCT TCC TCA AAT GCT GCT GAG CGA ACA CCC AGT CAC AGT	1774
His Ser Tyr Ala Ser Ser Asn Ala Ala Glu Arg Thr Pro Ser His Ser	
580 585 590	
GGT CAC CTC CAG GGT GAG CAT CCC TAC CTG ACA CCA TCC CCA GAG TCT	1822
Gly His Leu Gln Gly Glu His Pro Tyr Leu Thr Pro Ser Pro Glu Ser	
595 600 605	

FIG.11D

SUBSTITUTE SHEET (RULE 26)

38/68

CCT GAC CAG TGG TCA AGT TCA TCA CCC CAC TCT GCT TCT GAC TGG TCA	1870
Pro Asp Gln Trp Ser Ser Ser Ser Pro His Ser Ala Ser Asp Trp Ser	
610 615 620	
GAT GTG ACC ACC AGC CCT ACC CCT GGG GGT GCT GGA GGA GGT CAG CGG	1918
Asp Val Thr Thr Ser Pro Thr Pro Gly Gly Ala Gly Gly Gly Gln Arg	
625 630 635	
GGA CCT GGG ACA CAC ATG TCT GAG CCA CCA CAC AAC AAC ATG CAG GTT	1966
Gly Pro Gly Thr His Met Ser Glu Pro Pro His Asn Asn Met Gln Val	
640 645 650 655	
TAT GCG TGAGAGAGTC CACCTCCAGT GTAGAGACAT AACTGACTTT TGTAATGCT	2022
Tyr Ala	
GCTGAGGAAC AAATGAAGGT CATCCGGGAG AGAAATGAAG AAATCTCTGG AGCCAGCTTC	2082
TAGAGGTAGG AAAGAGAAGA TGTTCCTATT CAGATAATGC AAGAGAAGCA ATTCGTCAGT	2142
TTCACTGGGT ATCTGCAAGG CTTATTGATT ATTCTAATCT AATAAGACAA GTTTGTGGAA	2202
ATGCAAGATG AATACAAGCC TTGGGTCCAT GTTACTCTC TTCTATTTGG AGAATAAGAT	2262
GGATGCTTAT TGAAGCCCAG ACATTCTTGC AGCTTGGACT GCATTTTAAG CCCTGCAGGC	2322
TTCTGCCATA TCCATGAGAA GATTCTACAC TAGCGTCCTG TTGGGAATTA TGCCCTGGAA	2382
TTCTGCCCTGA ATTGACCTAC GCATCTCCTC CTCCTTGGAC ATTCTTTTGT CTTCAATTTGG	2442
TGCTTTTGGT TTTGCACCTC TCCGTGATTG TAGCCCTACC AGCATGTTAT AGGGCAAGAC	2502
CTTTGTGCTT TTGATCATTC TGGCCCATGA AAGCAACTTT GGTCTCCTTT CCCCTCCTGT	2562
CTTCCCGGTA TCCCTTGGAG TCTCACAAGG TTTACTTTGG TATGGTTCTC AGCACAAACC	2622
TTTCAAGTAT GTTGTCTCTT TGGAAAATGG ACATACTGTA TTGTGTTCTC CTGCATATAT	2682
CATTCCTGGA GAGAGAAGGG GAGAAGAATA CTTTCTTCA ACAAATTTTG GGGGCAGGAG	2742
ATCCCTTCAA GAGGCTGCAC CTTAATTTTT CTGTCTGTG TGCAGGTCTT CATATAAACT	2802

FIG.11E

SUBSTITUTE SHEET (RULE 26)

39/68

TTACCAGGAA GAAGGGTGTG AGTTTGTGT TTTTCTGTGT ATGGGCCTGG TCAGTGTA	2862
TTTATCCT TGATAGTCTA GTTACTATGA CCTCCCCAC TTTTAAAA CCAGAAAAAG	2922
GTTTGAATG TTGGAATGAC CAAGAGACAA GTTAACTCGT GCAAGAGCCA GTTACCCACC	2982
CACAGGTCCC CCTACTTCCT GCCAAGCATT CCATTGACTG CCTGTATGGA ACACATTGT	3042
CCCAGATCTG AGCATTCTAG GCCTGTTTCA CTCCTCACC CAGCATATGA AACTAGTCTT	3102
AACTGTTGAG CCTTTCCTTT CATATCCACA GAAGACACTG TCTCAAATGT TGTACCCTTG	3162
CCATTTAGGA CTGAACTTTC CTTAGCCCAA GGGACCCAGT GACAGTTGTC TTCCGTTTGT	3222
CAGATGATCA GTCTCTACTG ATTATCTTGC TGCTTAAAGG CCTGCTCACC AATCTTCTT	3282
TCACACCGTG TGGTCCGTGT TACTGGTATA CCCAGTATGT TCTCACTGAA GACATGGACT	3342
TTATATGTTT AAGTGCAGGA ATTGGAAAGT TGGACTTGTT TTCTATGATC CAAAACAGCC	3402
CTATAAGAAG GTTGGAAAAAG GAGGAACTAT ATAGCAGCCT TTGCTATTTT CTGCTACCAT	3462
TTCTTTTCCT CTGAAGCGGC CATGACATTC CCTTTGGCAA CTAACGTAGA AACTCAACAG	3522

FIG.11F

40/68

AACATTTTCC TTTCCTAGAG TCACCTTTTA GATGATAATG GACAACTATA GACTTGCTCA	3582
TTGTTTCAGAC TGATTGCCCC TCACCTGAAT CCACTCTCTG TATTCATGCT CTGGGCAATT	3642
TCTTTGACTT TCTTTTAAGG GCAGAAGCAT TTTAGTTAAT TGTAGATAAA GAATAGTTTT	3702
CTTCCTCTTC TCCTTGGGCC AGTTAATAAT TGGTCCATGG CTACACTGCA ACTTCCGTCC	3762
AGTGCTGTGA TGCCCATGAC ACCTGCAAAA TAAGTTCCTG CTGGGCATTT TGTAGATATT	3822
AACAGGTGAA TTCCCGACTC TTTTGGTTTG AATGACAGTT CTCATTCCTT CTATGGCTGC	3882
AAGTATGCAT CAGTGCTTCC CACTTACCTG ATTTGTCTGT CCGTGGCCCC ATATGGAAAC	3942
CCTGCGTGTG TGTTGGCATA ATAGTTTACA AATGGTTTTT TCAGTCCTAT CCAAATTTAT	4002
TGAACCAACA AAAATAATTA CTCTGCCCT GAGATAAGCA GATTAAGTTT GTTCATTCTC	4062
TGCTTTATTG TCTCCATGTG GCAACATTCT GTCAGCCTCT TTCATAGTGT GCAAACATTT	4122
TATCATTCTA AATGGTGACT CTCTGCCCTT GGACCCATTT ATTATTCACA GATGGGGAGA	4182
ACCTATCTGC ATGGACCCTC ACCATCCTCT GTGCAGCACA CACAGTGCAG GGAGCCAGTG	4242
GCGATGGCGA TGACTTTCTT CCCCTG	4268

FIG. 11G

41/68

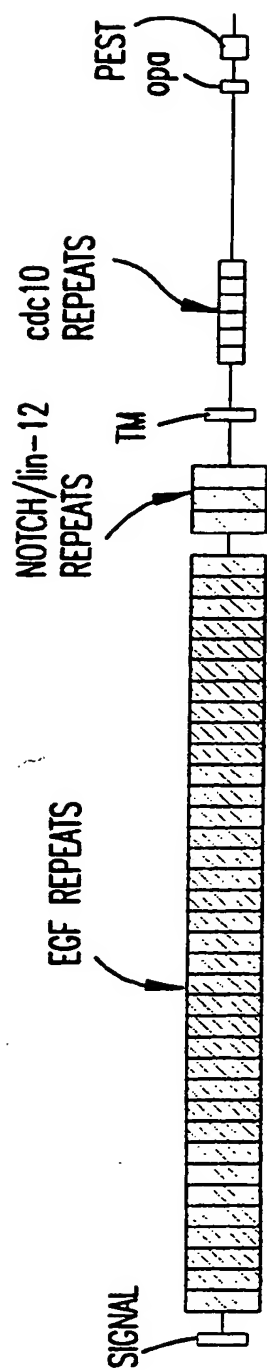
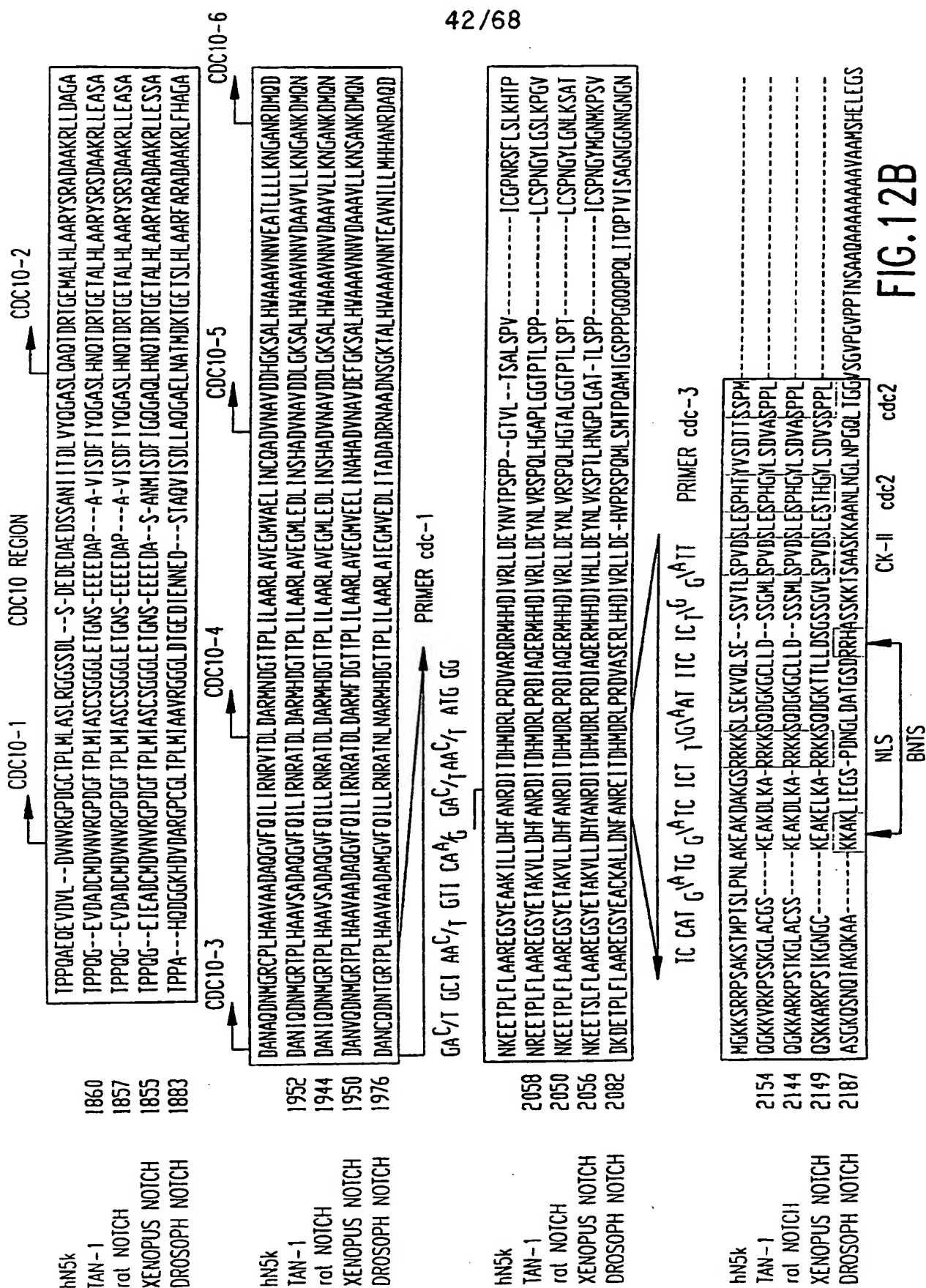


FIG.12A





## hNSK

## hN5k

NSK

NSN

## PEST-CONTAINING REGION

**FIG. 12C**

44/68

		Potential signal cleavage site							
hum N	MP	—	—	—	ALRPAL	LWALLALWLC	CA	—	APA HA
TAN-1	MP	—	—	—	PL	LAPLLCLALL	PA	—	LAA RG
Xen N	MD	—	—	—	—	RIGLAVLLCS	LP	—	VLT QG
Dros N	MQSQRSRRRS	RAPNTWICFW	INKMHAVASL	PASLPLLLLT	LAFANLPNIV	RGTDALVAA			
hum N	MLGKATCRCA	SGFTGEDCQY	STSHPCFVSR	PCLNGGTCHM	LSRDT-YECT	CQVGFTGKEC			
Tan-1	GVADYACSCA	LGFSGPLCLT	PLDNAC-LTN	PCRNGGTCOL	LT-LTEYKCR	CPPGWSGKSC			
Xen N	NAIDFICHCP	VGFTDKVCLT	PVDNAC-VNN	PCRNGGTCEL	LNSVTEYKCR	CPPGWTGDSC			
Dros N	GRPGISCKCP	LGFDLSCEI	AVPNAC-DHV	TCLNGGTCOL	KT-LEEYCA	CANGYTGERC			
hum N	NLPGSYQQQC	PQFTGQYCD	SLYVPCAPSP	CVNGGTCRQT	GDFTFECNCL	PGFEGSTCER			
TAN-1	NEVGSYRCVC	RATHGPNCE	RPYVPCSPSP	CQNGGTCRPT	GDVTHECACL	PGFTGQNCEE			
Xen N	NEFGSYRCTC	QNRFTGRNCD	EPYVPCNPSP	CLNGGTCRQT	DDTSYDCTCL	PGFSGQNCEE			
Dros N	NTHGSYQCMC	PTGYTGKDCD	TKYNPCSPSP	CQNAGICRSN	G-LSYECKCP	KGFEGKNCEQ			

## EGF-like Repeats

QCRDGYEPCV	NEGMCVITYHN	GTGYCKCEP	FLGEYQHRD	PCE-KNRCQN	GGTC-VAQA	83
RCSQPGETCL	NGGKCEA-AN	GTEACVCGGA	FVGPRCQDPN	PCL-STPCKN	AGTCHVDDR	80
RCTQTAEMCL	NGGRCEMTPG	GTGVCLCGNL	YFGERCQFPN	PCTIKNQCMN	FGTCEPVLG	90
SCTSVG-CQ	NGGTCVTQLN	GKTYCACDSH	YVGDYCEHRN	PCN-SMRCQN	GGTCQVTFRN	117
QWTDACLSHP	CANGSTCTTV	-ANQFSCKC	LTGFTGQKCE	TDVNEC-DIP	GHCQHGCTCL	199
QQADPCASNP	CANGGQCLPF	-EASYICHG	PPSFHGPTCR	QDVNECCQKP	RLCRHGGTCH	196
QQADPCASNP	CANGGKCLPF	-EIQYICKC	PPGFHGATCK	QDINEC-S-Q	NPCKNGGQCI	195
ETKNLCASSP	CRNGATCTAL	AGSSSFCTSC	PPGFTGDTCS	YDIEEC-Q-S	NPCKYGGICV	233
NIDDCPNHRC	QNGGVCVDGV	NTYNCRCPPO	WTGQFCTEDV	DECLLPNA-	CQNGGTCANR	318
NIDDCPGNNC	KNGGACVDGV	NTYNCPPE	WTGQYCTEDV	DECQLMPNA-	CQNGGTCHNT	315
NIDDCPSNNC	RNGGTCVDGV	NTYNQCAPPD	WTGQYCTEDV	DECQLMPNA-	CQNGGTCHNT	314
NYDDCLGHLG	QNGGTCIDGI	SDYTCRCPPN	FTGRFCQDDV	DECAQRDHPV	CQNGATCTNT	352

FIG.13A

45/68

hum N	INGGYGCVCVN	GWSGDDCSEN	IDDCAFASCT	PGSTCIDRVA	SFSQMCPEGK	AGLLCHLDDA
TAN-1	HGGYNCVCVN	GTGEDCSEN	IDDCASAACF	HGATCHORVA	SFYCECPHGR	TGLLCHLND
Xen N	YGGYNCVCVN	GTGEDCSEN	IDDCANAACH	SGATCHORVA	SFYCECPHGR	TGLLCHLDNA
Dros N	HGSYSCICVN	GWAGLDCSNN	TDDCKQAACF	YGATCIDGVG	SFYCQCTKGK	TGLLCHLDDA

hum N	AFHCECLKGY	AGPRCEMDIN	ECHSDPCQND	ATCLDKIGGF	TCLCMPGFKG	VHCELEINEC
TAN-1	SFECQCLQGY	TGPRCEIDVN	ECVSNPCQND	ATCLDQIGEF	QCMMPGYEG	VHCEVNTDEC
Xen N	SFQCNCPOGY	AGPRCEIDVN	ECLSNPCOND	STCLDQIGEF	QCICMPGYEG	LYCETNIDEC
Dros N	SYRCNC SQG	TGPRCETNIN	ECESHPCQNE	GSCLDDPGTF	RCVCMPGFTG	TQCEIDIDEC

hum N	ATGFTGVLCE	ENIDNCDPDP	CHHGQCQDGI	DSYTCICNPG	YMGATCSDQI	DECYSSPCLN
TAN-1	TEGYTGTHCE	VDIDCDPDP	CHYGSKDGV	ATFTCLCRPG	YTGHHCEINI	NECSSOPCRL
Xen N	TEGFTGRHCE	QDINECIPDP	CHYGTCKDGI	ATFTCLCRPG	YTGRLCDNDI	NECLSKPCLN
Dros N	PPGYTGTSCE	ININDCDSNP	CHRGKCIDDV	NSFKCLCDPG	YTG YICQKQI	NECESNPCQF

CISNPCHKGA	LCDTNPLNGQ	YICTCPQGYK	GADCTEDVDE	CAMANSNPCE	HAGKCVNTDG	438
CISNPCNEGS	NCDTNPVNGK	AICTCPSGYT	GPACSQDVDE	CSLG-ANPCE	HAGKCINTLG	434
CISNPCNEGS	NCDTNPVNGK	AICTCPPGYT	GPACNNDVDE	CSLG-ANPCE	HGGRCTNTLG	433
CTSNPCHADA	ICDTSPINGS	YACSCATGYK	GVDCESEDIDE	CDQG-SPCE	HNGICVNTPG	470

QSNPCVNNGQ	CVDKVNRFQC	LCPPGFTGPV	CQIDIDDCSS	TPCLNGAKCI	DHPNGYECQC	558
ASSPCLHNGR	CLDKINEFQC	ECPTGFTGHL	CQYDVDECAS	TPCKNGAKCL	DGPNTYTCVC	554
ASNPCLHNGK	CIDKINEFRC	DCPTGFSGNL	CQHDFDECTS	TPCKNGAKCL	DGPNSYTCQC	553
QSNPCLNDGT	CHDKINGFKC	SCALGFTGAR	CQINIDDCQS	QPCRNRGICH	DSIAGYSCEC	590

DGRCIDLUNG	YQCNCQPGTS	GVNCEINFDD	CASNPCIHG-	ICMDGINRYS	CVCSPGFTGQ	677
RGTCQDPDNA	YLCFCLKGTT	GPNCEINLDD	CASSPCDSG-	TCLDKIDGYE	CACEPGYTGS	673
GGQCTDRENG	YICTCPKGTT	GVNCETKIDD	CASNLCDNG-	KCIDKIDGYE	CTCEPGYTCK	672
DGHCQDRVGS	YYCQCQAGTS	GKNCEVNVNE	CHSNPCNNGA	TCIDGINSYK	CQCVPGFTGQ	710

FIG.13B

46 / 68

hum N	RCNIDIDECA	SNPCRKGATC	INGVNGFRCI	CPEGPHHPSC	YSQVNECLSN	PCI-HGNCTG
TAN-1	MCNSNIDECA	GNPCHNGGTC	EDGINGFTCR	CPEGYHDPTC	LSEVNECNSN	PCV-HGACRD
Xen N	LCNININECD	SNPCRNGGTC	KDQINGFTCV	CPDGYHDHMC	LSEVNECNSN	PCI-HGACHD
Dros N	HCEKNVDECI	SSPCANNGVC	IDQVNGYKCE	CPRGFYDAHC	LSDVDECASN	PCVNEGRCD
hum N	DECASNPCLN	QGTCFDDISG	YTCHCVLPYT	GKNCQTVLAP	CSPNPCENAA	VCKESPNFES
TAN-1	NECASNPCLN	KGTCIDDVAG	YKNCCLLPYT	GATCEVVLAP	CAPSPCRNGG	ECROSEDYES
Xen N	NECSSNPCLN	HGTCIDDVAG	YKNCMLPYT	GAICEAVLAP	CAGSPCKNGG	RCKESEDFT
Dros N	DDCVTNPCGN	GGTCIDKVN	YKCVCKVPFT	GRDCESKMDP	CASNRCKNEA	KCTPSSNFLD
hum N	CLANPCQNGG	SCMDGVNTFS	CLCLPGFTGD	KCQTDNMECL	SEPCKNGGTC	SDYVNSYTCK
TAN-1	CRPNPCHNGG	SCTDGINAF	CDCLPGFRGT	FCEEDINECA	SDPCRNGANC	TDCVDSYTCT
Xen N	CQPNPCHNGG	SCSDGINMFF	CNCPAGFRGP	KCEEDINECA	SNPCKNGANC	TDCVNSYTCT
Dros N	CASFPCQNGG	TCLDGIIDYS	CLCVDGFDGK	HCETDINECL	SQPCQNGATC	SOYVNSYTCT

GLSGYKCLCD	AGWVGINCEV	DKNECLSNPC	QNGGTCDNLV	NGYRCTCKKG	FKGYNCQVNI	796
SLNGYKDCDD	PGWSGTNCDI	NNNECESNPC	VNGGTCKDMT	SGIVCTCREG	FSGPNCQTNI	792
GVNGYKDCDE	AGWSGSNCDI	NNNECESNPC	MNGGTCKDMT	GAYICTCKAG	FSGPNCQTNI	791
GINEFICHCP	PGYTGRCEL	DIDECSSNPC	QHGGTCYDKL	NAFSCQCMGP	YTGQKCEINI	830

YTCLCA-PGW	QGQRCTIDID	EC-ISKPCMN	HGLCHINTQGS	YMCECPPGFS	GMDCEEDIDD	914
FSCVCPTAGA	KGQTCEVDIN	EC-VLSPCRH	GASCQNTGGS	YRCHCQAGYS	GRNCETDIDD	911
FSCECP-PGW	QGQTCEIDMN	EC-VNRPCRN	GATCQNTNGS	YKCNCKPGYT	GRNCMDIDD	909
FSCTCK-LGY	TGRYCDDEDID	ECSLSSPCRN	GASCLNVPGS	YRCLCTKGYE	GRDCAINTDD	949

COAGFDGVHC	ENNINECTES	SCFNGGTCVD	GINSFSLCP	VGFTGSFCLH	EINECSSHPC	1034
CPAGFSGIHC	ENNTPDCTES	SCFNGGTCVD	GINSFTCLCP	PGFTGSYCOH	VVNECDSPRC	1031
CQPGFSGIHC	ESNTPDCTES	SCFNGGTCID	GINTFTCQCP	PGFTGSYCOH	DINECDSPKC	1029
CPLGFSGINC	QTNDEDCTES	SCLNGGSCID	GINGYNCSC	AGYSGANCQY	KLKNCDSNPC	1069

FIG.13C

47/68

hum N	LNEGTCVDGL	GTYRCSCPLG	YTGKNCOTLV	NLCSRSPCKN	KGTCVQKAE	SQCLCPSGWA
TAN-1	LLGGTCQDGR	GLHRC TCPQG	YTGPNCONLV	HWCDSSPCKN	GGKCWQHTQ	YRCECPSGWT
Xen N	LNGGTCQDSY	GTYKCTCPQG	YTGLNCONLV	RWCDSSPCKN	GGKCWQTNNF	YRCECKSGWT
Dros N	LNGATCHEQN	NEYTCHCPSG	FTGKQCSEYV	DWCGQSPCEN	GATCSQMKHQ	FSCCKSAGWT

hum N	SNPCQHGATC	SDFIGGYRCE	CVPGYQGVNC	EYEVECCNQ	PCQNGGTCID	LVNHFKCSCP
TAN-1	PSPCQNGATC	TDYLGGSCK	CVAGYHGVNC	SEEIDECLSH	PCQNGGTCLD	LPNTYKCSGP
Xen N	PNPCQNGATC	TDYLGGSCE	CVAGYHGVNC	SEEINECLSH	PCQNGGTCID	LINTYKCSGP
Dros N	SQPCQNGGTC	RDIGAYECQ	CRQGFQGNQ	ELNIDDCAPN	PCQNGGTCHD	RVMNFSCSGP

hum N	CLSNPCSSEG	SLDCIQLTND	YLCVCRSAFT	GRHCETFDV	CPQMPCLNGG	TCAVASNMPD
TAN-1	CLSNPCDARG	TQNCVQRVND	FHCECRAGHT	GRRCESVING	CKGKPCKNNG	TCAVASNTAR
Xen N	CLSNPCDSRG	TQNCIQLVND	YRCECRQFT	GRRCESVVDG	CKGMPCRNGG	TCAVASNTAR
Dros N	CLSNPCSNAG	TLDCVQLVNN	YHCNCRPGHM	GRHCEHKVDF	CAQSPCQNGG	NCNI—RQS

GAYCDVPNVS	CDIAASRRGV	LVEHLCOHSG	VCINAGNTHY	CQCPLGYTGS	YCEEQLDECA	1154
GLYCDVPSVS	CEVAAQRQGV	DVARLCOHGG	LCVDAGNTHH	CRCQAGYTGS	YCEDLVDECS	1151
GVYCDVPSVS	CEVAAKQGV	DIVHLCRNSG	MCVDTGNTHF	CRCQAGYTGS	YCEEQVDECS	1149
GKLCQVQTIS	CQDAADRKGL	SLRQLC—NNG	TCKDYGNSHV	CYCSQGYAGS	YCQKEIDECQ	1188

PGTRGLLEE	NIDDCAR—	—GPHCLN	GGQCMDRIGG	YSCRCLPGFA	GERCEGDINE	1267
RGTQGVHCEI	NVDDCNPPVD	PVSRSPKCFN	NGTCVDQVGG	YSCTCPPGFV	GERCEGDVNE	1271
RGTQGVHCEI	NVDDCTPFYD	SFTLEPKCFN	NGKCIDRVGG	YNCICPPGFV	GERCEGDVNE	1269
PGTMGIICEI	NKDDCKP—	—GACHN	NGSCIDRVGG	FECVCQPGFV	GARCEGDINE	1300

GFICRCPPGF	SGARCQS—	SCGQVKCRKG	EQCVHTAS—	GPRFCPSP—	—RDCE—	1376
GFICKCPAGF	EGATCENDAR	TCGSLRCLNG	GTCISGPR—	SPTCLCLGPF	TGPECQFPAS	1389
GFICKCPGF	DGATCEYDSR	TCSNLRCONG	GTCISVLT—	SSKVCSEGY	TGATCQYPVI	1387
GHHICNNGF	YGKNCESGQ	DCDSNPCRVG	—NCVVADEGF	GYRCEPRGT	LGEHCEIDTL	1415

FIG.13D

48/68

hum N	-GC-ASSPCQ HGGSCHPQRQ PPYYSCQCAP PFSGSRCEL- -YTAPP- -S- -TPP
TAN-1	SPCLGGNPCY NQGTCEPTSE SPFYRCLCPA KFNGLLCHIL DYSFGG- -GAGRDIPPP
Xen N	SPC-ASHPCY NGGTCQFFAE EPFFQCFCPK NFNGLFCHIL DYEFG- -GLGKNITPP
Dros N	DEC-SPNPCA QGAACEDLLG D- YECLCPS KWKGRCDIY DANYPGWNGG SGSCNDRYAA

hum N	NN-QCDELN TVECLFDNFE CQGNSTCK- -YDKYCADHF KDNHCNQGNC SEECGWDGLD
TAN-1	SDGHCDQCN SAGCLFDGFD CQRAEGQCN LYDQYCKDH SDGHCDQCN SAECEWDGLD
Xen N	NDGKCDQCN NTGCLYDGF CQKVEQCN LYDQYCKDH QDGHCDQCN NAECEWDGLD
Dros N	KNGKCNEECN NAACHYDGH CERKLKSCDS LFDAYCQKH YDGFCDYCN NAECSWDGLD

hum N	YYGEKSAAMK KQ- R- - - - - MTRRSL PGEQ- - E QEVAGSKVFL
TAN-1	YYGEEELRK HPIKRAAEGW AAPDALLQV KASLLPGGSE GRRRRRELDP MDVRGSIVYL
Xen N	YYGNEELKK HHIKRSTDYW SDAPSAI- - FSTMESIL LGRHRELDE MEVRGSIVYL
Dros N	WKDNVRVPEI EDTDFARKNK ILYTQVHQ- - - - - TGIIYL

## LNR (Notch/Lin-12 Repeats)

-A- TCL	SOYCADKARD GVCDEACNSH ACQWDGGDCS LTMENPWANC SSPLPCWDYI	1476
LIEE- ACE	LPECQEDAGN KVCSLQCNH ACQWDGGDCS LNFNDPWKNC TQSLQCKYF	1501
DND- ICE	NEQCEADN KVCNANCNH ACQWDGGDCS LNFNDPWKNC TQSLQCKYF	1498
DLEQORAMCD	KRGCTEQGN GICSDCNTY ACNFDGNDCS LGI-NPWANC TAN-EXWKF	1531

CAADQPEN-L	AEGTLVIVL MPPEQLLODA R-SFLRALGT LLHTNLRKR DSQELMVYP	1591
CAEHVPER-L	AAGTL-VVV LMPPEQLRNS SFHFLRELSR VLHTNVVFKR DAHQQMIFP	1619
C-ANMPEN-L	AEGTLVLVVL MPPERLKNNS V-NFLRELSR VLHTNVVFKK DSKGEYKIYP	1615
CENKTQSPVL	AEGAMSVML MNVEAFREIQ A-QFLRMNSH MLRTTVRLKK DALGHDIIIN	1650

EIDNRQCVQD	SDHCFKNTDA AAALLASHAI QG- - TLSYP LVSWSSESLT PERT-Q-LLY	1680
EIDNRQCVQA	SSQCFQSATD VAAFLGALAS LGSL-NIPYK IEAVQSETVE PPPAQ-LHF	1737
EIDNRQCYKS	SSQCFNSATD VAAFLGALAS LGSLDLSYK IEAVKSENME TPKPST-LYP	1730
EIDNRKCTEC	FTHAVEAAEF LAATAAKHQL RNDQ-IHSV RGIKNPGDED NGEPPANVKY	1745

FIG.13E

49/68

hum N	LLAVAVVIL FIILLGVIMA KRKRK—HGS LWLPEGFTLR RDASNHKRRE PVGQDAVGLK
TAN-1	MYVAAAFLV LFFVGCGVLL SRKRRRQHGO LWFPEGFKV- SEASKKKRRE ELGEDSVGLK
Xen N	MLSMVLPLL IIFVFMVIV NKKRRREHDS FGSPTALFOK NPA-KRNGET PW-EDSVGLK
Dros N	VITGIIILVII ALAFFGMVL- STQRKRAHGV TWPEGFAP AAVMSRRRRD PHQEMRNLN

## CDC-10/Ankyrin Repeats

hum N	PIDRRPWTQQ HLEAADIRRT PSLALTPPQA EQEVDVLDVN VRGPDGCTPL MLASLRGGSS
TAN-1	QTDHRQWTQQ HLDAADL-RM SAMAPTPPQG EVDADCMDVN VRGPDGFTPL MIASCSGGGL
Xen N	KTDPRQWTRQ HLDAADL-RI SSMAPTPPQG EIEADCMDVN VRGPDGFTPL MIASCSGGGL
Dros N	EADQRVWSQA HLDVVDV-R- AIM—TPP-A HQDGGKHVDV ARGPCGLTPL MIAAVRGGGL

hum N	ANACDNMGRG PLHAAVAADA QGVFQILIRN RVTDLARMN DGTTPILAA RLAVEGMVAE
TAN-1	ANIQDNMGRG PLHAAVSADA QGVFQILIRN RATDLARMH DGTTPILAA RLAVEGMLED
Xen N	ANVQDNMGRG PLHAAVAADA QGVFQILIRN RATDLARMF DGTTPILAA RLAVEGMVEE
Dros N	ANCQDNTGRG PLHAAVAADA MGVFQILLRN RATNLNARMH DGTTPILAA RLAEIGMVED

NLSVOVSEAN LIGTGTEHW VDDE—	—G POPKKVKAED EALLSE-EDD	1782
PLK-NASDGA LMDDNQN—W GDED—	—LETKKRFEE PVVLPD-LDD	1837
PIK-NMTDGS FMDDNQN—W GDEET—	—LENKRFRFEE QVILPELVDD	1831
KQVAMQSQGV GQPGAH—W SDDSDMPLP KRQRSDPVSG	VGLGNNGGYA SDHTMVSEYE	1861

DLSDEDEDAE DSSANIITDL VYQGASLQAO TDRGTGEMALH LAARYSRADA AKRLLDAGAD	1902
ETGNSEEE-E DAPA-VISDF IYQGASLHNQ TDRGTGETALH LAARYSRADA AKRLLEASAD	1954
ETGNSEEE-E DASANMISDF IGQGAQLHNQ TDRGTGETALH LAARYARADA AKRLLESSAD	1949
DTGEDIENNE DSTAQVISDL LAQGAELNAT MDKTGETSLH LAARFARADA AKRLLDAGAD	1976

LINCQADVNA VDDHGKSALH WAAAVNNVEA TLLLLKNGAN RDMQDNKEET PLFLAAREGS	2022
LINSHADVNA VDDLKKSALH WAAAVNNVDA AVLLKNGAN KDMQNNREET PLFLAAREGS	2074
LINAHADVNA VDEFGKSALH WAAAVNNVDA AAVLLKNSAN KDMQNNKEET SLFLAAREGS	2069
LITADADINA ADNSGKTALH WAAAVNNTEA VNILLMHAN RDAQDDKDET PLFLAAREGS	2096

FIG.13F

50/68

hum N	Y E A A K I L L D H	F A N R D I T D H M	D R L P R D V A R D	R M H D I V R L L	D E Y N V T P S P P	— G T V L — T S
TAN-1	Y E T A K V L L D H	F A N R D I T D H M	D R L P R D I A Q E	R M H D I V R L L	D E Y N L V R S P Q	L H G A P L G G T P
Xen N	Y E T A K V L L D H	Y A N R D I T D H M	D R L P R D I A Q E	R M H D I V H L L	D E Y N L V K S P T	L H N G P L G A T —
Dros N	Y E A C K A L L D N	F A N R E I T D H M	D R L P R D V A S E	R L H D I V R L L	D E — H V P R S P Q	M L S M T P Q A M I

	NLS		CK II	cdc2	cdc2	
hum N	G S R R K K S L S E	K V Q L S E — S S	V T L S P V D S L E	S P H T Y V S D T T	S S P M	— — — — —
TAN-1	A — R R K K S Q D G	K G C L L D — S S	G M L S P V D S L E	S P H G Y L S D V A	S P P L	— — — — —
Xen N	A — R R K K S Q D G	K T T L L D S G S S	G V L S P V D S L E	S T H G Y L S D V S	S P P L	— — — — —
Dros N	G S — P D N G L D A	T G S L R R K A S S	K K T S A A S K K A	A N L N G L N P G Q	L T G G V S G V P G	V P P T N S A A Q A
	B N T S					

hum N	— — — — —	— — — — —	— — — — —	I T S P G I L Q A S	P N P M L — A T A	A P P A P V H A Q H
TAN-1	— — — — —	— — — — —	— — — — —	L P S P F — Q Q S	P S V P L N H L P G	M P D T H L G I G H
Xen N	— — — — —	— — — — —	— — — — —	M T S P F — Q Q S	P S M P L N H L T S	M P E S Q L G M N H
Dros N	Y E D C I K N A Q S	M Q S L Q N G L D	M I K L D N Y A Y S	M G S P F — Q Q E	L L N G Q L G M N	G N G Q R N G V G P
	CK II			cdc2		

ALSPV — — — — —	— — — — —	ICGP	N R S F L S L K H T	P M G K K S R R P S	A K S T M P T S L P	N L A K E A K D A K	2127
TLSP — — — — —	— — — — —	LCSP	N G Y L G S L K P G	V Q G K K V R K P S	S K G L A C G S —	— — — — —	2178
TLSP — — — — —	— — — — —	ICSP	N G Y M G N M K P S	V Q S K K A R K P S	I K G N G C — — —	— — — — —	2170
G S P P P Q Q Q P	Q L I T Q P T V I S		A G N G G N N G N G	N A S G K Q S N Q T	A K Q K A A — — —	— — — — —	2208

— — — — —	— — — — —	— — — — —	— — — — —	— — — — —	— — — — —	— — — — —	2169
— — — — —	— — — — —	— — — — —	— — — — —	— — — — —	— — — — —	— — — — —	2219
— — — — —	— — — — —	— — — — —	— — — — —	— — — — —	— — — — —	— — — — —	2213
AAAAAAVAA	M S H E L E G S P V	G V G M G G N L P S	P Y D T S S M Y S N	A M A A P L A N G N	P N T G A K Q P P S	— — — — —	2327

AL S F S N L H E M	Q — — — — —	— — — — —	— — — — —	— P L A H G A S T V	L P S V S Q L L S H	H H I V S P G S —	2235
L N V A A — K P E M	A A L G G G R L A	F E T G P P R L S H	L P V A S G T S T V	L G S S S G G A L N	F T V G G S T S L N	— — — — —	2306
I N M A T — K Q E M	A A — G S N R M A	F D A M V P R L T H	L — N A S S P N T I	M S — — — — —	N G S M H F T V G G A P T M N	— — — — —	2294
G V L P G G L C G M	G G L S G A G N G N	S H E Q G L S P P Y	S N Q S P P H S V Q	S S L A L S P H A Y	L G S P S P A K S R	— — — — —	2445

FIG. 13G



51/68

hum N GSAGSLRLH PVPVPADW— MNRMEVNETQ YNEMFGMVL PAEG-THPGI APOSRPPEGK  
 TAN-1 GQCEWLSRLQ SGMVNPQYNP LRGSVAPCPL STQAPSLQHG -MVGPLHSSL AASALSQMS  
 Xen N SQCDWLARLQ NGMVQNYDP IRNGIQGN- AQQAQALQHG LMTS-LHNGL PATTLSQMMT  
 Dros N PSLPTSPTHI QAMRHATQK QFGGSLNSL LGGANGGVV GGGGGGGGV GGGPQNSPV

hum N APQPSTCPP AVAGPLPTMY QIP—EM ARL-PSVAF TAMPQQDQ VAQTILPAYH  
 TAN-1 PPQPHLGVS AASGHLGRSF LSGEPSQADV QPLGPSSLAV HTILPQ-ESP ALPTSLPSSL  
 Xen N MQQQHHN-SS TTSTHINSF CSSDISQDQ QM—SSNI HSVMPQ-DTQ IFAASLPSNL  
 Dros N QQQLGGLEFG SAGLDLNG-F CGSPDSFHSG QMNPPS—I QSSMSG-SSP STNMLSPSSQ

hum N SDWSDVTTSP TGGAGGGQR GPGTHMSEPPHNN MQVYA  
 TAN-1 SDWSEGVSSP PT—SMQ SQIARIEAFK  
 Xen N SDWSEGISSP PT—SMQ PQRTHIEAFK  
 Dros N SDWSEGVQSP AANNLYISGG HQANKGSEAIYI

—HITPRE PLPP-IV-TF QLIPKGSIAQ PAG— 2320  
 —YQGLPSTRL ATQPHLVQTO QVQPNLQMQ QONLQPANIQ QQQSLQPPPP 2414  
 —YQAMPNTRL ANQPHLMQAO QMQQQN— —LQLHQS 2384  
 LGIISPTGSD MGIMLAPPQS SKNSAIMQTI SPQQQQQQQQ QQQQQHQQQQ QQQQQQQQQQ 2565

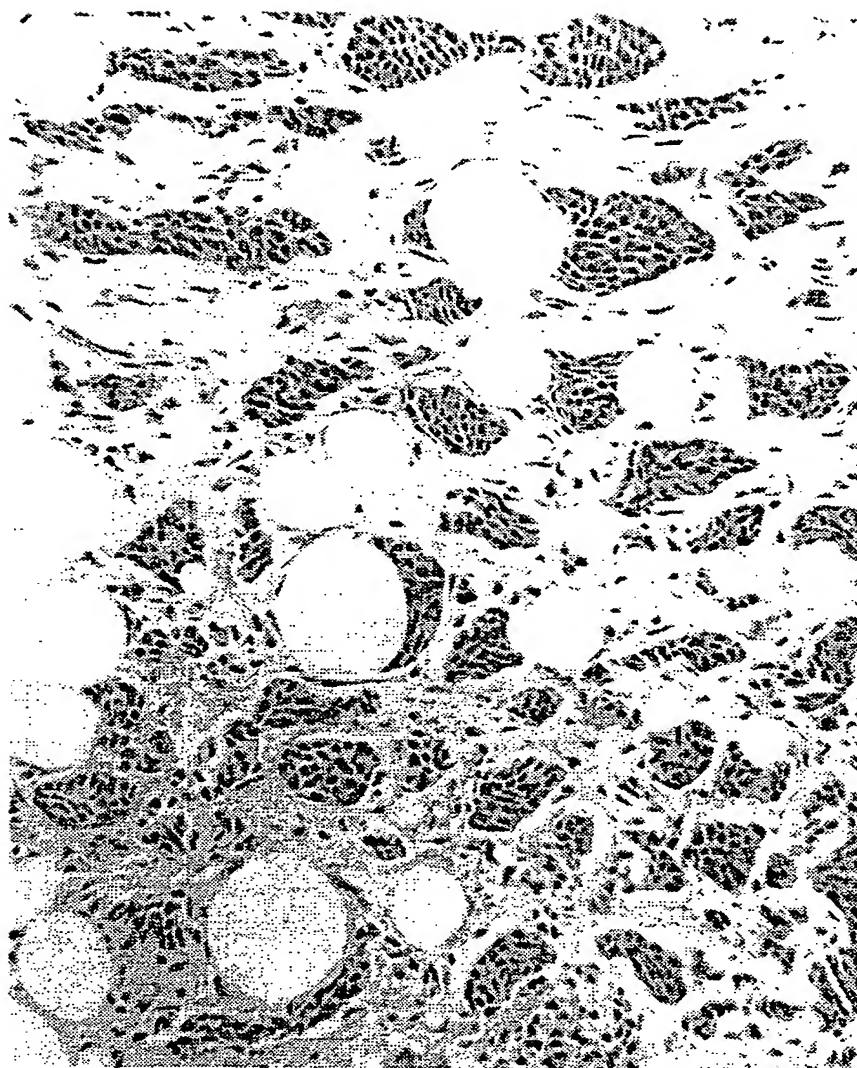
# PEST -containing Region

PFPASVGKYP TTPSQHSYAS SNAARTPSH SGHLOGEHPY LTPSPESPDQ WSSSSPHSA— 2433  
 VPPVTAAQFL TTPSQHSY-S S-PVENTPSH QLQVP-EGPF LTPSPESPDQ WSSSSPHSNV 2530  
 TQSMITAAQFL TTPSQHSY-S S-PMDNTPSH QLQVP-DHPF LTPSPESPDQ WSSSSPHSNM 2497  
 HNQQAFYQYL TTPSQHS— —GCHTPQH LVQTL-D-SY PTPSPESPGH WSSSSPRSN— 2671

2471  
 2556  
 2523  
 2703

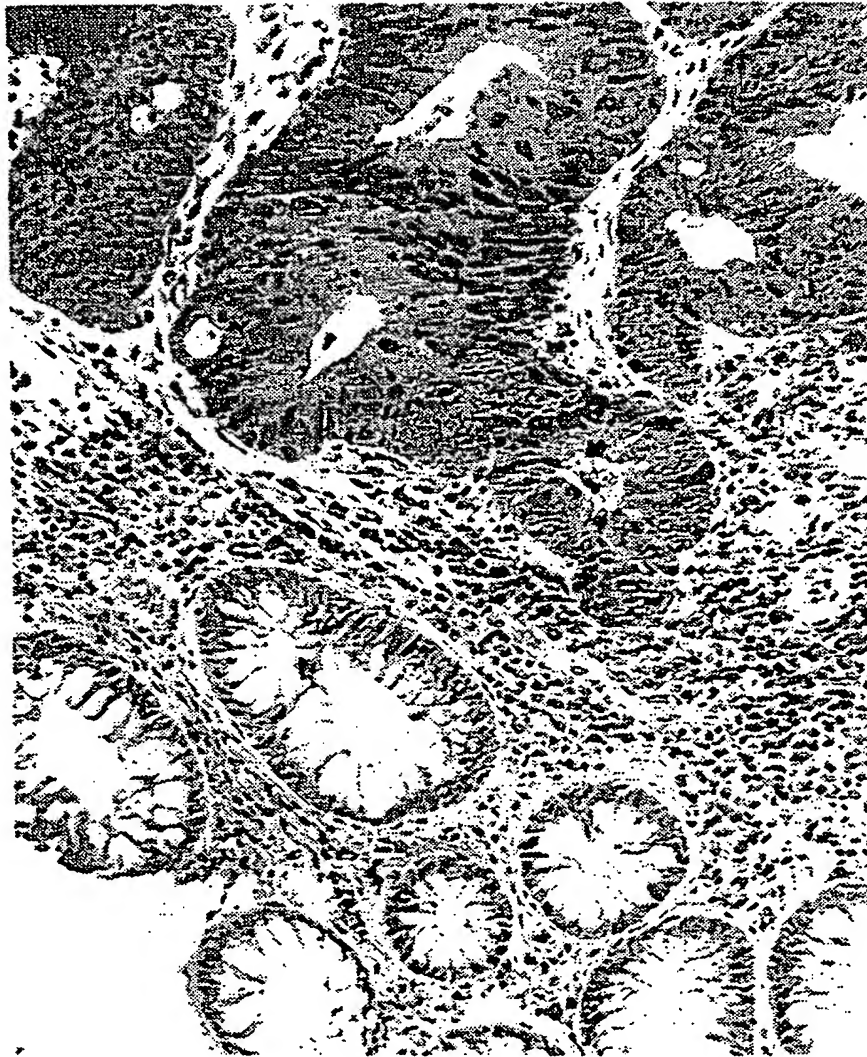
FIG.13H

52/68



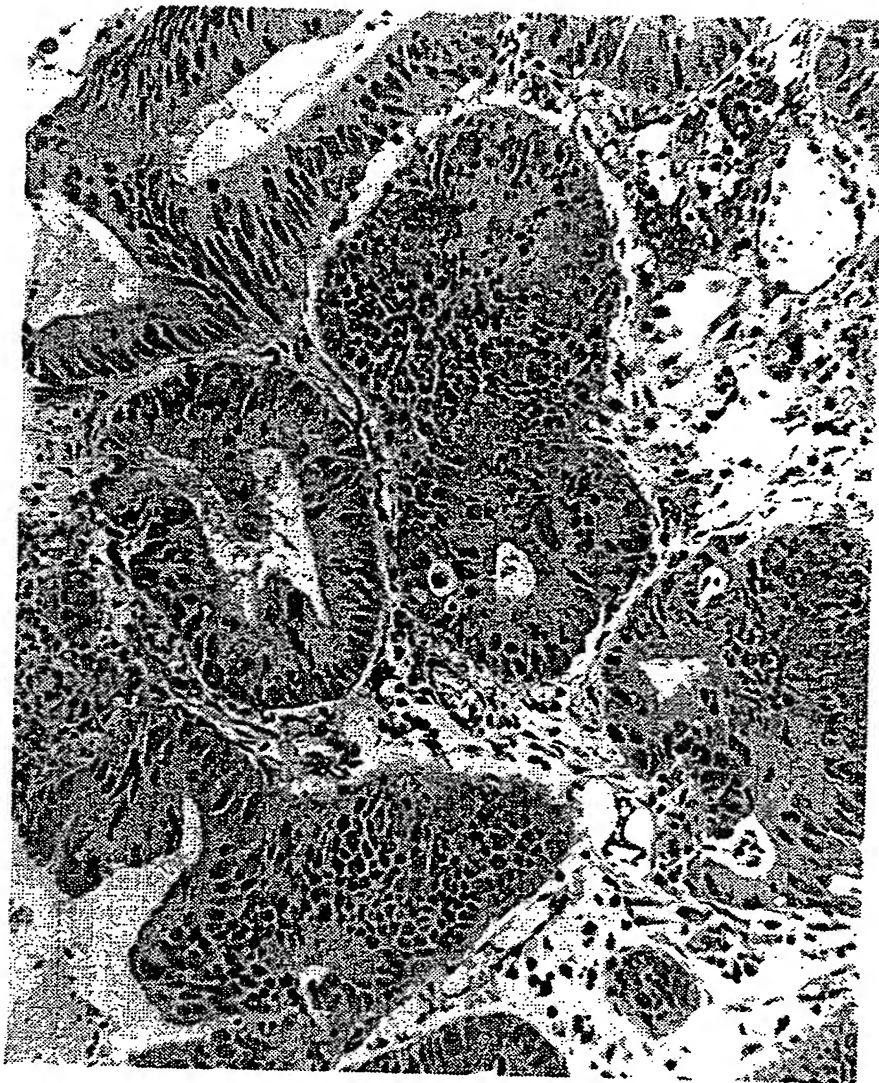
**FIG.14**

53/68



**FIG. 15A**

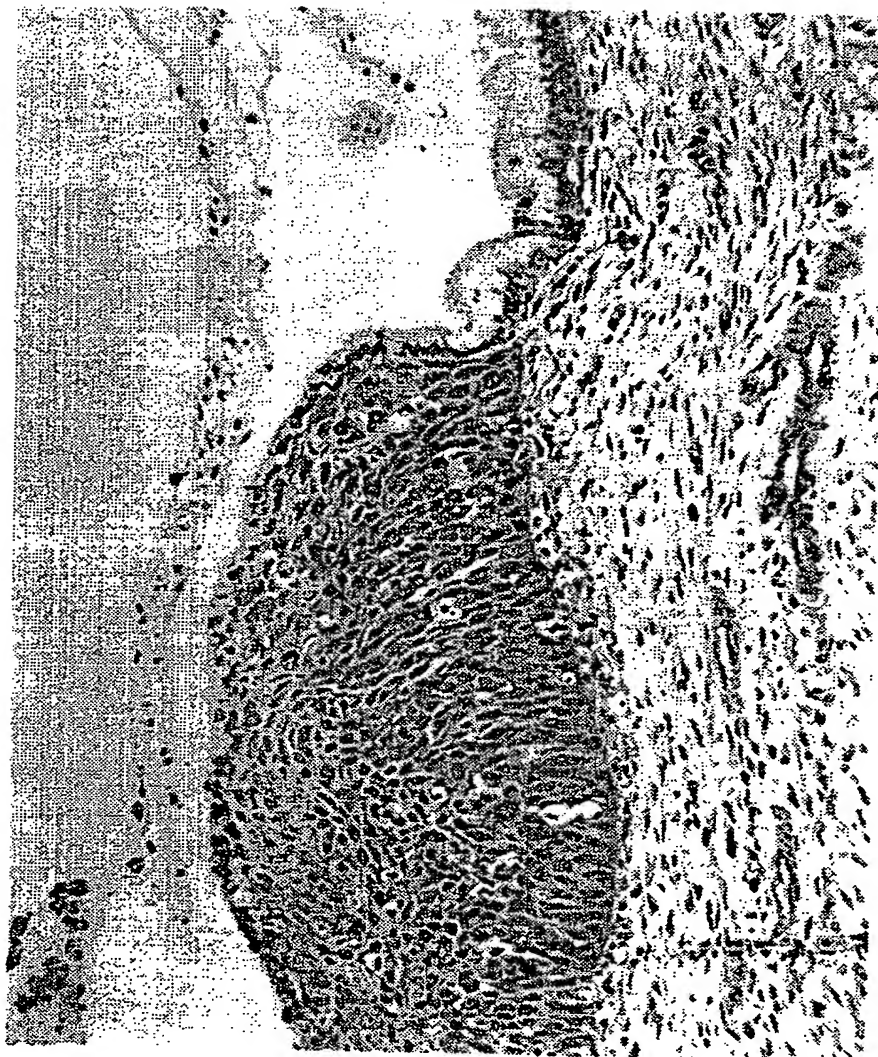
54/68



**FIG. 15B**

**SUBSTITUTE SHEET (RULE 26)**

55/68



**FIG. 16A**

**SUBSTITUTE SHEET (RULE 26)**

56/68



**FIG. 16B**



57/68

10 20 30 40 50 60 70 80 90  
\* \* \* \* \*  
GGAATTCCGC CCGCCCTGCG CCCCGCTCTG CTGTGGGGCG TGCTGGCGCT CTGGCTGTGC TCGCGGGCCC CCGCGCATGC ATTGCAGTGT  
P A L R P A L L W A L L A L W L C C A A P A H A L Q C>

100 110 120 130 140 150 160 170 180  
\* \* \* \* \*  
CGAGATGGCT ATGAACCCCTG TGTAAATGAA GGAATGTGTG TTACCTACCA CAATGGCACA GGATACTGCA AATGTCCAGA AGGCTTCTTG  
R D G Y E P C V N E G M C V T Y H N G T G Y C K C P E G F L>

190 200 210 220 230 240 250 260 270  
\* \* \* \* \*  
GGGGAATATT GTCAACATCG AGACCCCTGT GAGAAGAACC GCTGCCAGAA TGGTGGGACT TGTGTGGCCC AGGCCATGCT GGGGAAAGCC  
G E Y C Q H R D P C E K N R C Q N G G T C V A Q A M L G K A>

280 290 300 310 320 330 340 350 360  
\* \* \* \* \*  
ACGTGCCGAT GTGCCCTCAGG GTTACAGGA GAGGACTGCC AGTACTCAAC ATCTCATCCA TGCTTTGTGT CTCGACCCCTG CCTGAATGGC  
T C R C A S G F T G E D C Q Y S T S H P C F V S R P C L N G>

370 380 390 400 410 420 430 440 450  
\* \* \* \* \*  
GGCATATGCC ATATGCTCAG CCGGGATACC TATGAGTGCA CCTGTCAAGT CCGGTTTACA GGTAAGGAGT GCCAATGGAC GGATGCCCTGC  
G T C H M L S R D T Y E C T C Q V G F T G K E C Q W T D A C>

460 470 480 490 500 510 520 530 540  
\* \* \* \* \*  
CTGTCTCATC CCTGTGCAAA TGGAAGTACC TGTACCACTG TGGCCAACCA GTTCTCCTGC AAATGCCTCA CAGGCTTCAC AGGGCAGAAA  
L S H P C A N G S T C T T V A N Q F S C K C L T G F T G Q K>

550 560 570 580 590 600 610 620 630  
\* \* \* \* \*  
TGTGAGACTG ATGTCAATGA GTGTGACATT CCAGGACACT GCCAGCATGG TGGCACCTGC CTCAACCTGC CTGGTTCCTA CCAGTGCCAG  
C E T D V N E C D I P G H C Q H G G T C L N L P G S Y Q C Q>

640 650 660 670 680 690 700 710 720  
\* \* \* \* \*  
TGCCCTCAGG GCTTCACAGG CCACTACTGT GACAGCCTGT ATGTGCCCTG TGCACCTCA CCTTGTGTCA ATGGAGGCAC CTGTCCGCAG  
C P Q G F T G Q Y C D S L Y V P C A P S P C V N G G T C R Q>

730 740 750 760 770 780 790 800 810  
\* \* \* \* \*  
ACTGGTGA CTACTTTTGA GTGCAACTGC CTTCCAGGTT TTGAAGGGAG CACCTGTGAG AGGAATATTG ATGACTGCCC TAACCACAGG  
T G D F T F E C N C L P G F E G S T C E R N I D D C P N H R>

FIG.17A

SUBSTITUTE SHEET (RULE 26)

58/68

820 830 840 850 860 870 880 890 900  
\* \* \* \* \*  
TGTCAGAA TG GAGGGGTTG TGTGGATGG GTCAACACIT ACAACTGCCG CTGTCCCCCA CAATGGACAG GACAGTTCTG CACAGAGGAT  
C Q N G G V C V D G V N T Y N C R C P P Q W T G Q F C T E D>

910 920 930 940 950 960 970 980 990  
\* \* \* \* \*  
GTGGATGAAT GCCTGCTGCA GCCCAATGCC TGTCAAAATG GGGGCACCTG TGCCAACCGC AATGGAGGCT ATGGCTGTGT ATGTGTCAAC  
V D E C L L Q P N A C Q N G G T C A N R N G G Y G C V C V N>

1000 1010 1020 1030 1040 1050 1060 1070 1080  
\* \* \* \* \*  
GGCTGGAGTG GAGATGACTG CAGTGAGAAC ATTGATGATT GTGCCTTCGC CTCCTGTACT CCAGGCTCCA CCTGCATCGA CCGTGTGGCC  
G W S G D D C S E N I D D C A F A S C T P G S T C I D R V A>

1090 1100 1110 1120 1130 1140 1150 1160 1170  
\* \* \* \* \*  
TCCTTCTCTT GCATGTGCCC AGAGGGAAG GCAGGTCTCC TGTGTCATCT GGATGATGCA TGCATCAGCA ATCCTTGCCA CAAGGGGGCA  
S F S C M C P E G K A G L L C H L D D A C I S N P C H K G A>

1180 1190 1200 1210 1220 1230 1240 1250 1260  
\* \* \* \* \*  
CTGTGTGACA CCAACCCCT AAATGGGCAA TATATTGCA CCTGCCACAA AGGCTACAAA GGGGCTGACT GCACAGAAGA TGTGGATGAA  
L C D T N P L N G Q Y I C T C P Q G Y K G A D C T E D V D E>

1270 1280 1290 1300 1310 1320 1330 1340 1350  
\* \* \* \* \*  
TGTGCCATGG CCAATAGCAA TCCTTGTGAG CATGCAGGAA AATGTGTGAA CACGGATGGC GCCTTCCACT GTGAGTGTCT GAAGGGTTAT  
C A M A N S N P C E H A G K C V N T D G A F H C E C L K G Y>

1360 1370 1380 1390 1400 1410 1420 1430 1440  
\* \* \* \* \*  
GCAGGACCTC GTTGTGAGAT GGACATCAAT GAGTGCCATT CAGACCCCTG CCAGAATGAT GCTACCTGTC TGGATAAGAT TCGAGGCTTC  
A G P R C E M D I N E C H S D P C Q N D A T C L D K I G G F>

1450 1460 1470 1480 1490 1500 1510 1520 1530  
\* \* \* \* \*  
ACATGTCTGT GCATGCCAGG TTTCAAAGGT GTCCATTGTG AATTAGAAAT AAATGAATGT CAGAGCAACC CTTGTGTGAA CAATGGGCAG  
T C L C M P G F K G V H C E L E I N E C Q S N P C V N N G Q>

1540 1550 1560 1570 1580 1590 1600 1610 1620  
\* \* \* \* \*  
TGTGTGGATA AAGTCAATCG TTTCCAGTGC CTGTGCTTC CTGTTTCAC TGGGCCAGTT TGCCAGATTG ATATTGATGA CTGTTCCAGT  
C V D K V N R F Q C L C P P G F T G P V C Q I D I D D C S S>

FIG.17B



## 59/68

1630 1640 1650 1660 1670 1680 1690 1700 1710  
\* \* \* \* \*  
ACTCCGTGTC TGAATGGGCG AAAGTGTATC GATCACCCTGA ATGCTATGA ATGCCAGTGT GCCACAGGTT TCACTGGTGT GTTGTGTGAG  
T P C L N G A K C I D H P N G Y E C Q C A T G F T G V L C E>

1720 1730 1740 1750 1760 1770 1780 1790 1800  
\* \* \* \* \*  
GAGAACATTG ACAACTGTGA CCCCGATCCT TGCCACCATG GTCAGTGTCA GGATGGTATT GATTCCTACA CCTGCATCTG CAATCCCGGG  
E N I D N C D P D P C H H G Q C Q D G I D S Y T C I C N P G>

1810 1820 1830 1840 1850 1860 1870 1880 1890  
\* \* \* \* \*  
TACATGGGCG CCATCTGCAG TGACCAGATT GATGAATGTT ACAGCAGCCC TTGCCTGAAC GATGGTGGCT GCATTGACCT GGTCAATGGC  
Y M G A I C S D Q I D E C Y S S P C L N D G R C I D L V N G>

1900 1910 1920 1930 1940 1950 1960 1970 1980  
\* \* \* \* \*  
TACCAGTGCA ACTGCCAGCC AGGCACGTCA GGGTTAATT GTGAAATTA TTTTGATGAC TGTGCAAGTA ACCCTTGTAT CCATGGAATC  
Y Q C N C Q P G T S G V N C E I N F D D C A S N P C I H G I>

1990 2000 2010 2020 2030 2040 2050 2060 2070  
\* \* \* \* \*  
TGTATGGATG GCATTAATCG CTACAGTTGT GTCTGCTCAC CAGGATTCAC AGGCAGAGA TGTAAATTG ACATTGATGA GTGTGCCTCC  
C M D G I N R Y S C V C S P G F T G Q R C N I D I D E C A S>

2080 2090 2100 2110 2120 2130 2140 2150 2160  
\* \* \* \* \*  
AATCCCTGTC GCAAGGTGC AACATGTATC AACGGTGTGA ATGGTTTCCG CTGTATATGC CCCGAGGGAC CCCATACCC CAGCTGCTAC  
N P C R K G A T C I N G V N G F R C I C P E G P H H P S C Y>

2170 2180 2190 2200 2210 2220 2230 2240 2250  
\* \* \* \* \*  
TCACAGGTGA ACGAATGCCT GAGCAATCCC TGCATCCATG GAACTGTAC TGGAGGTCTC AGTGGATATA AGTGTCTCTG TGATGCAGGC  
S Q V N E C L S N P C I H G N C T G G L S G Y K C L C D A G>

2260 2270 2280 2290 2300 2310 2320 2330 2340  
\* \* \* \* \*  
TGGTTGGCA TCAACTGTGA AGTGGACAAA AATGAATGCC TTTGCAATCC ATGCCAGAAT GGAGGAATT GTGACAATCT GGTGAATGGA  
W V G I N C E V D K N E C L S N P C Q N G G T C D N L V N G>

2350 2360 2370 2380 2390 2400 2410 2420 2430  
\* \* \* \* \*  
TACAGGTGTA CTTGCAAGAA GGGCTTTAAA GGCTATAACT GCCAGGTGAA TATTGATGAA TGTGCCTCAA ATCCATGCCT GAACCAAGGA  
Y R C T C K F G F K G Y N C Q V N I D E C A S N P C L N Q G>

FIG.17C

60/68

2440 2450 2460 2470 2480 2490 2500 2510 2520  
\* \* \* \* \*  
ACCTGCTTTG ATGACATAAG TGGCTACACT TGCCACTGTG TGCTGCCATA CACAGGCAAG AATTGTCAGA CAGTATTGGC TCCCTGTTC  
T C F D D I S G Y T C H C V L P Y T G K N C Q T V L A P C S>

2530 2540 2550 2560 2570 2580 2590 2600 2610  
\* \* \* \* \*  
CCAAACCCCTT GTGAGAATGC TGCTGTTTGC AAAGAGTCAC CAAATTTTGA GAGTTATACT TGCTTGCTG C TCCTGGCTG GCAAGGTCAG  
P N P C E N A A V C K E S P N F E S Y T C L C A P G W Q G Q>

2620 2630 2640 2650 2660 2670 2680 2690 2700  
\* \* \* \* \*  
CGGTGTACCA TTGACATTGA CGAGTGTATC TCCAAGCCCT GCATGAACCA TGGTCTCTGC CATAACACCC AGGGCAGCTA CATGTGTGAA  
R C T I D I D E C I S K P C M N H G L C H N T Q G S Y M C E>

2710 2720 2730 2740 2750 2760 2770 2780 2790  
\* \* \* \* \*  
TGTCCACCAG GCTTCAGTGG TATGGACTGT GAGGAGGACA TTGATGACTG CCTTGCCAAT CCTTGCCAGA ATGGAGGTT C TGATGGA  
C P P G F S G M D C E E D I D D C L A N P C Q N G G S C M D>

2800 2810 2820 2830 2840 2850 2860 2870 2880  
\* \* \* \* \*  
GGAGTGAATA CTTTCTCCTG CCTCTGCCCT CCGGGTTTCA CTGGGGATAA GTGCCAGACA GACATGAATG AGTGTCTGAG TGAACCCCTGT  
G V N T F S C L C L P G F T G D K C Q T D M N E C L S E P C>

2890 2900 2910 2920 2930 2940 2950 2960 2970  
\* \* \* \* \*  
AAGAATGGAG GGACCTGCTC TGA CTACGTC AACAGTTACA CTGCAAGTG CCAGGCAGGA TTTGATGGAG TCCATTGTGA GAACAACATC  
K N G G T C S D Y V N S Y T C K C Q A G F D G V H C E N N I>

2980 2990 3000 3010 3020 3030 3040 3050 3060  
\* \* \* \* \*  
AATGAGTGCA CTGAGAGCTC CTGTTTCAAT GGTGGCACAT GTGTTGATGG GATTAACCTC TTCTCTTGCT TGTGCCCTGT GGGTTTCACT  
N E C T E S S C F N G G T C V D G I N S F S C L C P V G F T>

3070 3080 3090 3100 3110 3120 3130 3140 3150  
\* \* \* \* \*  
GGATCCTTCT GCCTCCATGA GATCAATGAA TGCAGCTCTC ATCCATGCC T GAATGAGGGA ACGTGTGTTG ATGGCCTGGG TACCTACCC  
G S F C L H E I N E C S S H P C L N E G T C V D G L G T Y R>

3160 3170 3180 3190 3200 3210 3220 3230 3240  
\* \* \* \* \*  
TGCAGCTGCC CCCTGGGCTA CACTGGGAAA AACTGTCAGA CCCTGGTGAA TCTCTGCAGT CCGTCTCCAT GTAAAAACAA AGGTACTTGT  
C S C P L G Y T G K N C Q T L V N L C S R S P C K N K G T C>

FIG.17D

SUBSTITUTE SHEET (RULE 26)

61/68

3250 3260 3270 3280 3290 3300 3310 3320 3330  
\* \* \* \* \*  
GTT CAG AAAA AAG CAG AGT C CAG TGC CTA TGT CCA TCT G GAT GGG CTG TGC CTA TTG T GAC GTG CCA ATG TCT CTTG TGAC ATAGCA  
V Q K K A E S Q C L C P S G W A G A Y C D V P N V S C D I A>

3340 3350 3360 3370 3380 3390 3400 3410 3420  
\* \* \* \* \*  
GCCTCCAGGA GAGGTGTGCT TGTGTAACAC TTGTGCCAGC ACTCAGGTGT CTGCATCAAT GCTGGCAACA CGCATTACTG TCAGTGCCCC  
A S R R G V L V E H L C Q H S G V C I N A G N T H Y C Q C P>

3430 3440 3450 3460 3470 3480 3490 3500 3510  
\* \* \* \* \*  
CTGGCCTATA CTGGGAGCTA CTGTGAGGAG CAACTCGATG AGTGTGGCTC CAACCCCTGC CAGCACGGGG CAACATGCAG TGACTTCATT  
L G Y T G S Y C E E Q L D E C A S N P C Q H G A T C S D F I>

3520 3530 3540 3550 3560 3570 3580 3590 3600  
\* \* \* \* \*  
GGTGGATACA GATGCCAGTG TGTCCAGGC TATCAGGTG TCAACTGTGA GTATGAAGTG GATGAGTGCC AGAATCAGCC CTGCCAGAAT  
G G Y R C E C V P G Y Q G V N C E Y E V D E C Q N Q P C Q N>

3610 3620 3630 3640 3650 3660 3670 3680 3690  
\* \* \* \* \*  
GGAGGCACCT GTATTGACCT TGTGAACCAT TTCAAGTGCT CTTGCCCACC AGGCACTCGG GGCTACTCT GTGAAGAGAA CATTGATGAC  
G G T C I D L V N H F K C S C P P G T R G L L C E E N I D D>

3700 3710 3720 3730 3740 3750 3760 3770 3780  
\* \* \* \* \*  
TGTGCCCCGG GTCCCCATTG CCTTAATGCT GGTCACTGCA TGGATAGGAT TGGAGCTAC AGTTGTGGCT GCTTGCTCGG CTTTGCTGGG  
C A R G P H C L N G G Q C M D R I G G Y S C R C L P G F A G>

3790 3800 3810 3820 3830 3840 3850 3860 3870  
\* \* \* \* \*  
GAGCGTTGTG AGGGAGACAT CAACGAGTGC CTCTCCAACC CCTGCAGCTC TGAGGGCAGC CTGGACTGTA TACAGCTCAC CAATGACTAC  
E R C E G D I N E C L S N P C S S E G S L D C I Q L T N D Y>

3880 3890 3900 3910 3920 3930 3940 3950 3960  
\* \* \* \* \*  
CTGTGTGTTT GCGTAGTGC CTTTACTGGC CGGCACTGTC AAACCTTCGT CGATGTGTGT CCCCAGATGC CCTGCCCTGAA TGGAGGGACT  
L C V C R S A F T G R H C E T F V D V C P Q M P C L N G G T>

3970 3980 3990 4000 4010 4020 4030 4040 4050  
\* \* \* \* \*  
TGTGCTGTGG CCACTAACAT GCCTGATGCT TTCAATTTGCC GTTGTCCTCC GGGATTTTCC GGGGCAAGGT GCCAGAGCAG CTGTGGACAA  
C A V A S N M P D G F I C R C P P G F S G A R C Q S S C G Q>

FIG.17E

SUBSTITUTE SHEET (RULE 26)

62/68

4060 4070 4080 4090 4100 4110 4120 4130 4140  
\* \* \* \* \*  
GTGAAATGTA GGAAGGGGA GCAGTGTGTG CACACCGCCT CTGGACCCCG CTGCTTCTGC CCCAGTCCCC GGGACTGCGA GTCAGGCTGT  
V K C R K G E Q C V H T A S G P R C F C P S P R D C E S G C>

4150 4160 4170 4180 4190 4200 4210 4220 4230  
\* \* \* \* \*  
GCCAGTAGCC CCTGCCAGCA CGGGGGCAGC TGCCACCCTC AGCGCCAGCC TCCTTATTAC TCCTGCCAGT GTGCCCCACC ATTCTCGGT  
A S S P C Q H G G S C H P Q R Q P P Y Y S C Q C A P P F S G>

4240 4250 4260 4270 4280 4290 4300 4310 4320  
\* \* \* \* \*  
AGCCGCTGTG AACTCTACAC GGCACCCCC AGCACCCCTC CTGCCACCTG TCTGAGCCAG TATTGTCCCG ACAAGCTCG GGATGGCGTC  
S R C E L Y T A P P S T P P A T C L S Q Y C A D K A R D G V>

4330 4340 4350 4360 4370 4380 4390 4400 4410  
\* \* \* \* \*  
TGTGATGAGG CCTGCAACAG CCATGCCTGC CAGTGGGATG GGGGTGACTG TTCTCTCACC ATGGAAGAAC CCTGGGCCAA CTGCTCCTCC  
C D E A C N S H A C Q W D G G D C S L T M E N P W A N C S S>

4420 4430 4440 4450 4460 4470 4480 4490 4500  
\* \* \* \* \*  
CCACTTCCTT GCTGGGATTA TATCAACAAC CAGTGTGATG AGCTGTGCAA CACGGTCGAG TGCCTGTTG ACAACTTTGA ATGCCAGGG  
P L P C W D Y I N N Q C D E L C N T V E C L F D N F E C Q G>

4510 4520 4530 4540 4550 4560 4570 4580 4590  
\* \* \* \* \*  
AACAGCAAGA CATGCAAGTA TGACAAATAC TGTGCAGACC ACTTCAAAGA CAACCACTGT AACCAGGGGT GCAACAGTGA GGAGTGTGGT  
N S K T C K Y D K Y C A D H F K D N H C N Q G C N S E E C G>

4600 4610 4620 4630 4640 4650 4660 4670 4680  
\* \* \* \* \*  
TGGGATGGGC TGGACTGTGC TGCTGACCAA CCTGAGAACC TGGCAGAAGG TACCCTGGTT ATTGTGGTAT TGATGCCACC TGAACAATG  
W D G L D C A A D Q P E N L A E G T L V I V V L M P P E Q L>

4690 4700 4710 4720 4730 4740 4750 4760 4770  
\* \* \* \* \*  
CTCCAGGATG CTGCAGCTT CTGCGGGCA CTGGGTACCC TGCTCCACAC CAACCTGGCG ATTAAGCGGG ACTCCCAGGG GGAATCATG  
L Q D A R S F L R A L G T L L H T N L R I K R D S Q G E L M>

4780 4790 4800 4810 4820 4830 4840 4850 4860  
\* \* \* \* \*  
GTGTACCCCT ATTATGGTGA GAAGTCAGCT GCTATGAAGA AACAGAGGAT GACACGCAGA TCCCTTCCTG GTGAACAAGA ACAGGAGGT  
V Y P Y Y G E K S A A M K K Q R M T R R S L P G E Q E Q E V>

FIG.17F

63/68

4870 4880 4890 4900 4910 4920 4930 4940 4950  
\* \* \* \* \*  
GCTGGCTCTA AAGTCTTTCT GGAAATTGAC AACCGCCAGT GTGTTCAAGA CTCAGACCAC TGCTTCAAGA ACACGGATGC AGCAGCAGCT  
A G S K V F L E I D N R Q C V Q D S D H C F K N T D A A A A>

4960 4970 4980 4990 5000 5010 5020 5030 5040  
\* \* \* \* \*  
CTCCTGGCCT CTCACGCCAT ACAGGGGACC CTGTCATACC CTCTTGTC TGTCGTCAGT GAATCCCTGA CTCAGAAGC CACTCAGCTC  
L L A S H A I Q G T L S Y P L V S V V S E S L T P E R T Q L>

5050 5060 5070 5080 5090 5100 5110 5120 5130  
\* \* \* \* \*  
CTCTATCTCC TTGCTGTTGC TGTGTGCATC ATTCTGTTA TTATCTGCT GGGGTAATC ATGGCAAAC GAAAGCGTAA GCATGGCTCT  
L Y L L A V A V V I I L F I I L L G V I M A K R K R K H G S>

5140 5150 5160 5170 5180 5190 5200 5210 5220  
\* \* \* \* \*  
CTCTGGCTGC CTGAAGTTT CACTCTTCCG CGAGATGCAA GCAATCACA GCGTCGTGAG CCAGTGGGAC AGGATGCTGT GGGGCTGAAA  
L W L P E G F T L R R D A S N H K R R E P V G Q D A V G L K>

5230 5240 5250 5260 5270 5280 5290 5300 5310  
\* \* \* \* \*  
AATCTCTCAG TGCAAGTCTC AGAAGCTAAC CTAATTGCTA CTGGAACAAG TGAACACTGG GTCGATGATG AAGGGCCCCA GCCAAAGAAA  
N L S V Q V S E A N L I G T G T S E H W V D D E G P Q P K K>

5320 5330 5340 5350 5360 5370 5380 5390 5400  
\* \* \* \* \*  
GTAAAGGCTG AAGATGAGGC CTTACTCTCA GAAGAAGATG ACCCATTTGA TCGACGGCCA TGGACACAGC AGCACCTTGA AGCTGCAGAC  
V K A E D E A L L S E E D D P I D R R P W T Q Q H L E A A D>

5410 5420 5430 5440 5450 5460 5470 5480 5490  
\* \* \* \* \*  
ATCCGTAGGA CACCATCGCT GGCTCTCACC CCTCCTCAGG CAGAGCAGGA GGTGGATGTC TTAGATGTGA ATGTCGCTGG CCCAGATGCC  
I R R T P S L A L T P P Q A E Q E V D V L D V N V R G P D G>

5500 5510 5520 5530 5540 5550 5560 5570 5580  
\* \* \* \* \*  
TGCACCCCAT TGATGTTGGC TTCTCTCCGA GGAGGCAGCT CAGATTGAG TGATGAAGAT GAAGATGCAG AGGACTCTTC TGCTAACATC  
C T P L M L A S L R G G S S D L S D E D E D A E D S S A N I>

5590 5600 5610 5620 5630 5640 5650 5660 5670  
\* \* \* \* \*  
ATCACAGACT TGGTCTACCA GGGTCCAGC CTCCAGGCCC AGACAGACCG GACTGGTGAG ATGGCCCTGC ACCTTGCAGC CCGCTACTCA  
I T D L V Y Q G A S L Q A Q T D R T G E M A L H L A A R Y S>

FIG.17G

SUBSTITUTE SHEET (RULE 26)

64/68

5680 5690 5700 5710 5720 5730 5740 5750 5760  
\* \* \* \* \*  
CGGGCTGATG CTGCCAAGCG TCTCCTGGAT GCAGGTGCAG ATGCCAATGC CCAGGACAAC ATGGGCCGCT GTCCACTCCA TGCTGCAGTG  
R A D A A K R L L D A G A D A N A Q D N M G R C P L H A A V>

5770 5780 5790 5800 5810 5820 5830 5840 5850  
\* \* \* \* \*  
GCAGCTGATG CCCAAGGTGT CTTCAGATT CTGATTGCA ACCGAGTAAC TGATCTAGAT GCCAGGATGA ATGATGGTAC TACACCCCTG  
A A D A Q G V F Q I L I R N R V T D L D A R M N D G T T P L>

5860 5870 5880 5890 5900 5910 5920 5930 5940  
\* \* \* \* \*  
ATCCTGGCTG CCCGCTGGC TGTCGAGGGA ATGGTGGCAG AACTGATCAA CTGCCAAGCG GATGTGAATG CAGTGGATGA CCATGGAAAA  
I L A A R L A V E G M V A E L I N C Q A D V N A V D D H G K>

5950 5960 5970 5980 5990 6000 6010 6020 6030  
\* \* \* \* \*  
TCTGCTCTTC ACTGGGCAGC TGCTGTCAAT AATGTGGAGG CAACTCTTTT GTTGTGAAA AATGGGGCCA ACCGAGACAT GCAGGACAAC  
S A L H W A A A V N N V E A T L L L L K N G A N R D M Q D N>

6040 6050 6060 6070 6080 6090 6100 6110 6120  
\* \* \* \* \*  
AAGGAAGAGA CACCTCTGTT TCTTGTGCC CGGGAGGGGA GCTATGAAGC AGCCAAGATC CTGTTAGACC ATTTTGCCAA TCGAGACATC  
K E E T P L F L A A R E G S Y E A A K I L L D H F A N R D I>

6130 6140 6150 6160 6170 6180 6190 6200 6210  
\* \* \* \* \*  
ACAGACCATA TGGATGCTCT TCCCCGGAT GTGGCTCGGG ATCGCATGCA CCATGACATT GTGCCCCTTC TGGATGAATA CAATGTGACC  
T D H M D R L P R D V A R D R M H H D I V R L L D E Y N V T>

6220 6230 6240 6250 6260 6270 6280 6290 6300  
\* \* \* \* \*  
CCAAGCCCTC CAGGCACCGT GTTGAATTCT GCTCTCTCAC CTGTCACTCG TGGGCCCAAC AGATCTTTCC TCAGCCTGAA GCACACCCCA  
P S P P G T V L T S A L S P V I C G P N R S F L S L K H T P>

6310 6320 6340 6350 6360 6370 6380 6390 6400  
\* \* \* \* \*  
ATGGGCAAGA AGTCTAGACG GCCCAGTGCC AAGAGTACCA TGCCTACTAG CCTCCCTAAC CTTGCCAAGG AGGCAAAGGA TGCCAAGGGT  
M G K K S R R P S A K S T M P T S L P N L A K E A K D A K G>

6400 6410 6420 6430 6440 6450 6460 6470 6480  
\* \* \* \* \*  
AGTAGGAGGA AGAAGTCTCT GAGTGAGAAG GTCCAATGT CTGAGAGTTC AGTAACTTTA TCCCCTGTTG ATTCCCTAGA ATCTCCTCAC  
S R R K K S L S E K V Q L S E S S V T L S P V D S L E S P H>

FIG.17H

SUBSTITUTE SHEET (RULE 26)

65/68

6490 6500 6510 6520 6530 6540 6550 6560 6570  
\* \* \* \* \*  
ACGTATGTTT CCGACACCAC ATCCTCTCCA ATGATTACAT CCCCTGGGAT CTTACAGGCC TCACCCAACC CTATGTTGGC CACTGCCGCC  
T Y V S D T T S S P M I T S P G I L Q A S P N P M L A T A A>  
6580 6590 6600 6610 6620 6630 6640 6650 6660  
\* \* \* \* \*  
CCTCCTGCCC CAGTCCATGC CCAGCATGCA CTATCTTTT CTAACCTTCA TGAATGCAG CCTTTGGCAG ATGGGGCCAG CACTGTGCTT  
P P A P V H A Q H A L S F S N L H E M Q P L A H G A S T V L>  
6670 6680 6690 6700 6710 6720 6730 6740 6750  
\* \* \* \* \*  
CCCTCAGTGA GCCAGTTGCT ATCCCACCAC CACATTGTGT CTCAGGCAG TGGCAGTGT GGAAGCTTGA GTAGGCTCCA TCCAGTCCCA  
P S V S Q L L S H H H I V S P G S G S A G S L S R L H P V P>  
6760 6770 6780 6790 6800 6810 6820 6830 6840  
\* \* \* \* \*  
GTCCAGCAG ATTGGATGAA CCGCATGGAG GTGAATGAGA CCCAGTACAA TGAGATGTTT GGTATGGTCC TGGCTCCAGC TGAGGGCACC  
V P A D W M N R M E V N E T Q Y N E M F G M V L A P A E G T>  
6850 6860 6870 6880 6890 6900 6910 6920 6930  
\* \* \* \* \*  
CATCCTGGCA TAGCTCCCCA GAGCAGGCCA CCTGAAGGGA AGCACATAAC CACCCCTCGG GAGCCCTTGC CCCCCATTGT GACTTTCAG  
H P G I A P Q S R P P E G K H I T T P R E P L P P I V T F Q>  
6940 6950 6960 6970 6980 6990 7000 7010 7020  
\* \* \* \* \*  
CTCATCCCTA AAGCAGTAT TGCCCAACCA GCGGGGGCTC CCCAGCCTCA GTCCACCTGC CCTCCAGCTG TTGCGGGCCC CCTGCCACC  
L I P K G S I A Q P A G A P Q P Q S T C P P A V A G P L P T>  
7030 7040 7050 7060 7070 7080 7090 7100 7110  
\* \* \* \* \*  
ATGTACCAGA TTCCAGAAAT GGCCCGTTTG CCCAGTGTGC CTTTCCCAC TGCCATGATG CCCCAGCAGG ACGGGCAGGT AGCTCAGACC  
M Y Q I P E M A R L P S V A F P T A M M P Q Q D G Q V A Q T>  
7120 7130 7140 7150 7160 7170 7180 7190 7200  
\* \* \* \* \*  
ATTCTCCCAG CCTATCATCC TTTCCCAGCC TCTGTGGGCA AGTACCCAC ACCCCCTTCA CAGCACAGTT ATGCTTCTC AAATGCTGCT  
I L P A Y H P F P A S V G K Y P T P P S Q H S Y A S S N A A>  
7210 7220 7230 7240 7250 7260 7270 7280 7290  
\* \* \* \* \*  
GAGCGAACAC CCAGTCACAG TGGTCACCTC CAGGGTGAGC ATCCCTACCT GACACCATCC CCAGAGTCTC CTGACCACTG GTCAAGTTCA  
E R T P S H S G H L Q G E H P Y L T P S P E S P D Q W S S S>

FIG.171

SUBSTITUTE SHEET (RULE 26)

66/68

7300	7310	7320	7330	7340	7350	7360	7370	7380
*	*	*	*	*	*	*	*	*
TCACCCACT	CTGCTTCTGA	CTGGTCAGAT	GTGACCACCA	GCCCTACCCC	TGGGGGTGCT	GGAGGAGGTC	AGCGGGGACC	TGGGACACAC
S P H	S A S D	W S D	V T T	S P T P	G G A	G G G	Q R G P	G T H>
7390	7400	7410	7420	7430	7440	7450	7460	7470
*	*	*	*	*	*	*	*	*
ATGTCTGAGC	CACCACACAA	CAACATGCAG	GTTTATGCGT	GAGAGAGTCC	ACCTCCAGTG	TAGAGACATA	ACTGACTTTT	GTAATGCTG
M S E	P P H N	N M Q	V Y A>					
7480	7490	7500	7510	7520	7530	7540	7550	7560
*	*	*	*	*	*	*	*	*
CTGAGGAACA	AATGAAGGTC	ATCCGGGAGA	GAAATGAAGA	AATCTCTGGA	GCCAGCTTCT	AGAGGTAGGA	AAGAGAAGAT	GTTCTTATTC
7570	7580	7590	7600	7610	7620	7630	7640	7650
*	*	*	*	*	*	*	*	*
AGATAATGCA	AGAGAAGCAA	TTGTCAGTT	TCACTGGGTA	TCTGCAAGGC	TTATTGATTA	TTCTAATCTA	ATAAGACAAG	TTTGTGGAAG
7660	7670	7680	7690	7700	7710	7720	7730	7740
*	*	*	*	*	*	*	*	*
TGCAAGATGA	ATACAAGCCT	TGGGTCCATG	TTTACTCTCT	TCTATTGGA	GAATAAGATG	GATGCTTATT	GAAGCCCAGA	CATTCTTGCA
7750	7760	7770	7780	7790	7800	7810	7820	7830
*	*	*	*	*	*	*	*	*
GCTTGACTG	CATTTTAAGC	CCTGCAGGCT	TCTGCCATAT	CCATGAGAAG	ATTCTACACT	AGCGTCCTGT	TGGGAATTAT	GCCCTGGAAT
7840	7850	7860	7870	7880	7890	7900	7910	7920
*	*	*	*	*	*	*	*	*
TCTGCCGAA	TTGACCTACG	CATCTCCTCC	TCCTTGACA	TTCTTTTGTC	TTCAATTGGT	GCTTTTGGTT	TTGCACCTCT	CCGTGATTGT
7930	7940	7950	7960	7970	7980	7990	8000	8010
*	*	*	*	*	*	*	*	*
AGCCCTACCA	GCATGTTATA	GGGCAAGACC	TTTGTGCTTT	TGATCATTTCT	GGCCCATCAA	AGCAACTTTG	GTCTCCTTTC	CCCTCCTGTC
8020	8030	8040	8050	8060	8070	8080	8090	8100
*	*	*	*	*	*	*	*	*
TTCCCGGTAT	CCCTTGAGT	CTCACAAGGT	TTACTTTGGT	ATGGTTCTCA	GCACAAACCT	TTCAAGTATG	TTGTTTCTTT	GGAAAATGGA
8110	8120	8130	8140	8150	8160	8170	8180	8190
*	*	*	*	*	*	*	*	*
CATACTGTAT	TGTGTTCTCC	TGCATATATC	ATTCTCGGAG	AGAGAAGGGG	AGAAGAATAC	TTTTCTTCAA	CAAATTTTGG	GGGCAGGAGA
8200	8210	8220	8230	8240	8250	8260	8270	8280
*	*	*	*	*	*	*	*	*
TCCCTTCAAG	AGGCTGCACC	TTAATTTTTC	TTGCTGTGT	GCAGGTCTTC	ATATAAACTT	TACCAGGAAG	AAGGGTGTA	GTTTGTGTT

FIG.17J

SUBSTITUTE SHEET (RULE 26)



67/68

8290	8300	8310	8320	8330	8340	8350	8360	8370
*	*	*	*	*	*	*	*	*
TTTCTGTGTA TGGCCCTGGT CAGTGTAAG TTTTATCCTT GATAGTCTAG TTAATATGAC CCTCCCCACT TTTTAAAAAC CAGAAAAAGC								
8380	8390	8400	8410	8420	8430	8440	8450	8460
*	*	*	*	*	*	*	*	*
TTTGAATGT TGAATGACC AAGAGACAAG TTAACCTGTC CAAGAGCCAG TTACCCACCC ACAGGTCCCC CTACTTCCTG CCAAGCATT								
8470	8480	8490	8500	8510	8520	8530	8540	8550
*	*	*	*	*	*	*	*	*
CATTGACTGC CTGTATGGAA CACATTGTC CCAGATCTGA GCATTCTAGG CCTGTTTCAC TCACTCACCC AGCATATGAA ACTAGTCTTA								
8560	8570	8580	8590	8600	8610	8620	8630	8640
*	*	*	*	*	*	*	*	*
ACTGTTGAGC CTTTCCTTC ATATCCACAG AAGACACTGT CTCAAATGTT GTACCCITGC CATTIAGGAC TGAACITTC TTAGCCCAAG								
8650	8660	8670	8680	8690	8700	8710	8720	8730
*	*	*	*	*	*	*	*	*
GGACCCAGTG ACAGTTGTCT TCCGTTTGC AGATGATCAG TCTCTACTGA TTATCTTGCT GCITAAAGGC CTGCTACCA ATCTTTCTT								
8740	8750	8760	8770	8780	8790	8800	8810	8820
*	*	*	*	*	*	*	*	*
CACACCGTGT GGTCCGTGTT ACTGGTATAC CCAGTATGTT CTCCTGAAG ACATGGACTT TATATGTCA AGTCAGGAA TTGGAAGTT								
8830	8840	8850	8860	8870	8880	8890	8900	8910
*	*	*	*	*	*	*	*	*
GGACTTGTTT TCTATGATCC AAAACAGCCC TATAAGAAGG TTGAAAAGG AGGAAGTATA TAGCAGCCTT TGCTATTTTC TGCTACCATT								
8920	8930	8940	8950	8960	8970	8980	8990	9000
*	*	*	*	*	*	*	*	*
TCTTTCTC TGAAGCGGCC ATGACATTCC CTTTGGCAAC TAAGTAGAA ACTCAACAGA ACATTTCTT TTCTAGAGT CACCTTTTAC								
9010	9020	9030	9040	9050	9060	9070	9080	9090
*	*	*	*	*	*	*	*	*
ATGATAATGG ACAACTATAG ACTTGCTCAT TGTCAGACT GATTGCCCT CACCTGAATC CACTCTCTGT ATTATGCTC TTGGCAATT								
9100	9110	9120	9130	9140	9150	9160	9170	9180
*	*	*	*	*	*	*	*	*
CTTTGACITTT CTTTAAAGG CAGAAGCATT TTAGTAAAT GTAGATAAG AATAGTITTC TTCTCTTCT CCTTGGCCA GTTAATAATT								
9190	9200	9210	9220	9230	9240	9250	9260	9270
*	*	*	*	*	*	*	*	*
GGTCCATGGC TAACTGCAA CTTCCGTCCA GTGCTGTGAT GCCCATGACA CCTGCAAAAT AAGTTCTGCC TGGGCATTTT GTAGATATTA								

FIG.17K

68/68

9280	9290	9300	9310	9320	9330	9340	9350	9360
*	*	*	*	*	*	*	*	*
ACAGGTGAAT	TCCCGACTCT	TTTGGTTTGA	ATGACAGTTC	TCATTCCTTC	TATGGCTGCA	AGTATGCATC	AGTGCTTCCC	ACTTACCTGA
9370	9380	9390	9400	9410	9420	9430	9440	9450
*	*	*	*	*	*	*	*	*
TTTGCTGTC	GGTGGCCCCA	TATGGAAACC	CTGCGTGTCT	GTTGGCATAA	TAGTTTACAA	ATGGTTTTTT	CAGTCCTATC	CAAATTTATT
9460	9470	9480	9490	9500	9510	9520	9530	9540
*	*	*	*	*	*	*	*	*
GAACCAACAA	AAATÁATTAC	TTCTGCCCTG	AGATAAGCAG	ATTAAGTTTG	TTCATTCTCT	GCTTTATTCT	CTCCATGTGG	CAACATTCTG
9550	9560	9570	9580	9590	9600	9610	9620	9630
*	*	*	*	*	*	*	*	*
TCAGCCTCTT	TCATAGTGTG	CAAACATTTT	ATCATTCTAA	ATGGTGACTC	TCTGCCCTTG	GACCCATTTA	TTATTCACAG	ATGGGGAGAA
9640	9650	9660	9670	9680	9690	9700	9710	9720
*	*	*	*	*	*	*	*	*
CCTATCTGCA	TGGACCTCA	CCATCCTCTG	TGCAGCACAC	ACAGTGCAGG	GAGCCAGTGG	CGATGGCGAT	GACTTTCTTC	CCCTGGGAAT

TCC

FIG.17L

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US93/09338

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) : Please See Extra Sheet.

US CL : 424/85.8; 435/6; 514/1, 2, 24

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/85.8; 435/6; 514/1, 2, 24

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 5,115,096 (SHOYAB ET AL) 19 MAY 1992, see column 8, lines 12-27, column 11, lines 39-63, column 14, lines 16-68, columns 16-18, column 28, Table IV, and Figure 13.	1-18, 22-28, 31, 34-40, 42-52, 55-58, 63-67, 75-94
A	US, A, 5,132,212 (KIRSCH ET AL) 21 JULY 1992, see column 7, lines 25-36.	23-31, 45, 55-60, 63-67, 84-90
A,E	US, A, 5,264,557 (SALOMON ET AL) 23 NOVEMBER 1993, see column 1, lines 20-49.	1-18, 22-28, 31, 34-40, 42-52, 55-58, 63-67, 75-94

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be part of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

13 December 1993

Date of signing of the international search report

29 DEC 1993

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile N . NOT APPLICABLE

Authorized officer

STEPHEN WALSH

Telephone No. (703) 308-0196

*S. Walsh for*

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US93/09338

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	THE NEW BIOLOGIST, Volume 2, No.7, issued July 1990, R.J. Greenspan, "The <u>Notch</u> Gene, Adhesion, And Developmental Fate In The <u>Drosophila</u> Embryo", pages 595-600, see abstract.	1-18,22-28,31,34-40,42-52,55-58,63-67,75-94
A	CELL, Volume 67, issued 15 November 1991, I. Rebay et al, "Specific EGF Repeats of Notch Mediate Interactions with Delta and Serrate: Implications for Notch as a Multifunctional Receptor", pages 687-699, see entire document.	1-94
A,P	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, Volume 90, issued May 1993, J.F. De Celis et al, "Genetic and molecular characterization of a Notch mutation in its Delta- and Serrate-binding domain in <u>Drosophila</u> ", pages 4037-4031, see page 4037, column 2, paragraph 2.	21,54
A	EUROPEAN JOURNAL OF BIOCHEMISTRY, Volume 190, issued May 1990, J.A. Campos-Ortega et al, "Molecular analysis of a cellular decision during embryonic development of <u>Drosophila melanogaster</u> : epidermogenesis or neurogenesis", pages 1-10, see section bridging pages 4-5, page 8, column 2, paragraphs 2 and 4.	1-20,22,34-40,42-44,46-53,75-83,91-94
A	Biological Abstracts, Volume 93, No.11, issued 01 June 1992, J. Robbins et al, "Mouse mammary tumor gene int-3: A member of the notch gene family transforms mammary epithelial cells", see page AB-465, abstract no. 122736, J. Virol., 66(4), 2594-2599.	23-28,31,45,55-58,63-67,84-90
A	CELL, Volume 66, issued 23 August 1991, L.W. Ellisen et al, "TAN-1, the Human Homolog of the <u>Drosophila Notch</u> gene, Is Broken by Chromosomal Translocations in T Lymphoblastic Neoplasms", pages 649-661, see pages 657-658.	1-18,22-28,31,34-40,42-52,55-58,63-67,75-94

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US93/09338

## A. CLASSIFICATION OF SUBJECT MATTER:

IPC (5):

A61K 31/00, 31/70, 37/02, 39/44, 39/395; C07H 21/04; G01N 33/53, 33/68

## B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

APS, BIOSIS, CA, INPADOC, JICST-E, MEDLINE, search terms: notch protein or gene product, delta protein or gene product, serrate protein or gene product, disease, disorder, cancer, DNA, nucleic acid, anti-sense, therapy or treatment or pharmaceutical

## BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

- I. Claims 1-18, 22, 34-40, 42-44, 46-52, 75-83 and 91-94, drawn to pharmaceutical compositions comprising a Notch protein, fragments, chimeras, derivatives or analogs of a Notch protein, methods of treating or preventing malignancy or nervous system disorder, a method of promoting tissue regeneration or repair, and a method of treating a benign dysproliferative disorder, classified in Class 514, subclass 2.
- II. Claims 19, 20 and 53, drawn to a pharmaceutical composition comprising a derivative or analog of a Delta protein and a method of treating or preventing a malignancy, classified in Class 514, subclass 2.
- III. Claims 21 and 54, drawn to a pharmaceutical composition comprising a derivative or analog of a Serrate protein, classified in Class 514, subclass 2.
- IV. Claims 23-28, 31, 45, 55-58, 63-67 and 84-90, drawn to a pharmaceutical composition comprising a nucleic acid encoding a Notch protein, fragments or chimeras of a Notch protein, a method of treating or preventing malignancy comprising administration of nucleic acid encoding a Notch protein, a method of treating a patient with a tumor, and a pharmaceutical composition comprising an isolated oligonucleotide consisting of at least six nucleotides and a recombinant cell, classified in Class 514, subclass 44.
- V. Claims 29 and 59, drawn to a pharmaceutical composition comprising nucleic acid encoding a fragment of a Delta protein, and a method of treating or preventing malignancy comprising administration of nucleic acid encoding a Delta protein, classified in Class 514, subclass 44.
- VI. Claims 30 and 60, drawn to a pharmaceutical composition comprising nucleic acid encoding a fragment of a Serrate protein, and a method of treating or preventing malignancy comprising administration of nucleic acid encoding a Serrate protein, classified in Class 514, subclass 44.
- VII. Claims 32, 33, 41, 61 and 62, drawn to a pharmaceutical composition comprising an antibody and a method of treating or preventing malignancy comprising administration of antibody, classified in Class 424, subclass 85.8.
- VIII. Claim 68-74, drawn to a method of diagnosing a disease, classified in Class 435, subclass 6.

